

AD\_\_\_\_\_

Award Number: W81XWH-04-C-0002

TITLE: Diagnostic Methods for Predicting Performance Impairment  
Associated with Combat Stress

PRINCIPAL INVESTIGATOR: Gerald Matthews, Ph.D.  
Joel S. Warm, Ph.D.  
David Washburn, Ph.D.

CONTRACTING ORGANIZATION: University of Cincinnati  
Cincinnati, Ohio 45221-0627

REPORT DATE: December 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.



# REPORT DOCUMENTATION PAGE

Form Approved  
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

<b>1. AGENCY USE ONLY</b> (Leave blank)		<b>2. REPORT DATE</b> December 2004	<b>3. REPORT TYPE AND DATES COVERED</b> Annual (1 Dec 2003 - 30 Nov 2004)	
<b>4. TITLE AND SUBTITLE</b> Diagnostic Methods for Predicting Performance Impairment Associated with Combat Stress			<b>5. FUNDING NUMBERS</b> W81XWH-04-C-0002	
<b>6. AUTHOR(S)</b> Gerald Matthews, Ph.D. Joel S. Warm, Ph.D. David Washburn, Ph.D.			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> University of Cincinnati Cincinnati, Ohio 45221-0627  <i>E-Mail:</i> matthegd@email.uc.edu				
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			<b>10. SPONSORING / MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b> Original contains color plates: All DTIC reproductions will be in black and white.				
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for Public Release; Distribution Unlimited				<b>12b. DISTRIBUTION CODE</b>
<b>13. ABSTRACT (Maximum 200 Words)</b>  This report reviews the first year of research on the diagnostic utility of psychophysiological indices that may predict the current and future functional efficiency of the soldier. The research focuses especially on the measurement of cerebral bloodflow using transcranial doppler sonography, together with additional indices including salivary cortisol and subjective stress state. Two studies, with one ongoing, have been conducted at the University of Cincinnati. The first study investigated cerebral bloodflow during a vigilance task requiring detection of an absent stimulus element, a task relevant to military surveillance. The second study developed a simulator method for investigating attention to hazards during vehicle driving, and demonstrated that subjective measures correlate with performance. The third study compares cerebral bloodflow, salivary cortisol and subjective states as diagnostic predictors of sensory vigilance. Preliminary results confirm reliable measurement of individual differences in bloodflow, but suggest that the bloodflow - performance association may be sensitive to moderator factors. Pilot work at Georgia State University has developed and validated a 'shoot/don't shoot' rifle marksmanship task and a test battery for assessment of individual differences in attentional functioning for use in subsequent studies. Implications of work completed for the future progress of the research are discussed.				
<b>14. SUBJECT TERMS</b> Stress, cerebral bloodflow, performance, vigilance, sustained attention, fatigue, cortisol, individual differences, workload				<b>15. NUMBER OF PAGES</b> 79
				<b>16. PRICE CODE</b>
<b>17. SECURITY CLASSIFICATION OF REPORT</b> Unclassified	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> Unclassified	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b> Unclassified	<b>20. LIMITATION OF ABSTRACT</b> Unlimited	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)  
Prescribed by ANSI Std. Z39-18  
298-102



## Table of Contents

Cover	01
SF 298	02
Introduction	04
Body	05
Key Research Accomplishments	67
Reportable Outcomes	68
Conclusions	69
References	72



## INTRODUCTION

Combat stress often disrupts performance efficiency and situation awareness during military operations. The assessment of fitness for duty in stressful environments requires techniques for monitoring the current functional efficiency of the soldier, along with techniques that will predict when a soldier will not be able to sustain performance on some future task. This report describes the accomplishments of the first year of a proposed three-year program of research that aims to investigate the diagnostic utility of psychophysiological indices that may predict concurrent and future functional efficiency. The research focuses especially on the measurement of cerebral blood flow using transcranial Doppler sonography, together with additional indices including salivary cortisol and subjective stress state. Two studies have been completed at the University of Cincinnati and a third is ongoing. The first study established that cerebral blood flow is sensitive to the workload of a vigilance task requiring detection of an absent stimulus element, corresponding to the military task of monitoring for the disappearance from surveillance of an enemy unit. The second study developed a simulator method for investigating attention to hazards during vehicle driving, and demonstrated that subjective measures correlate with performance. The third study is designed to compare cerebral blood flow, salivary cortisol, and subjective states as diagnostic predictors of sensory vigilance. Preliminary results confirm reliable measurement of individual differences in blood flow, but suggest that the blood flow – performance association may be sensitive to moderator factors. Pilot work at Georgia State University has developed and validated a 'shoot/don't shoot' rifle marksmanship task and a test battery for assessment of individual differences in attentional functioning for use in full-scale studies to be initiated early in the next year of the project. Implications of work completed for the future progress and timeline of the research are discussed.



## BODY OF REPORT

### Overview

Broadly, the aim for this project is to identify useful diagnostic predictors of the person's fitness to sustain performance of various, attentionally demanding tasks. A predictor of special interest is the cerebral blood flow index obtained through transcranial Doppler sonography (TCD). The project also evaluates the diagnosticity of subjective state, salivary cortisol, and performance of short, high workload tasks. The Statement of Work for this project specifies two overarching goals: (1) assessment of whether the various psychophysiological indices predict subsequent sustained performance, and (2) whether TCD provides a concurrent index of performance and functional efficiency during sustained performance. The three studies initiated in the first year of the project address these goals. Study 1 constituted a further test of the utility of the TCD as a concurrent index of operator status during a task requiring vigilance. Study 2 investigated the utility of subjective state as a correlate of a more complex sustained performance task - vehicle driving. Study 3 (in progress) tests the utility of the complete battery of psychophysiological indices of interest as a predictor of vigilance. A summary of the main findings and implications of these studies follows.

*Study 1.* It is well-known that observers experience difficulty in sustaining attention, evidenced by deterioration in signal detection during prolonged work (the *vigilance decrement*). Our earlier TCD studies (e.g., Mayleben, 1998, Hitchcock et al., 2003) demonstrated that the vigilance decrement is accompanied by declining cerebral blood flow, especially in the right hemisphere. Consistent with the workload model of vigilance (Warm & Dember, 1998), it appears that, like loss of perceptual sensitivity, cerebral blood flow is depressed only when the detection task imposes a high workload on the operator. These findings suggested that measurement of blood flow during sustained performance may provide a useful diagnostic index of the operator's continuing fitness to perform demanding signal detection tasks.

Most vigilance tasks, including those used in our prior research, require operators to detect a designated target stimulus. For example, a sentry might be required to detect the appearance of an enemy soldier (Johnson, 2000). However, in military settings, detection of the *absence* of a stimulus may be equally important. For example, an observer monitoring a formation of enemy tanks should notice if one of the tanks is missing from the formation. The aim of this study was to test the generality of the workload model to a task requiring detection of a missing stimulus element. A 40-minute signal detection task was employed. Observers ( $N = 96$ ) were required to search for either the presence or for the absence of a critical feature. It was expected that blood flow would show a greater decline when the target stimulus was defined by absence of a feature, a search which is believed to be more attentionally demanding than searching for feature presence (Treisman & Gormican, 1988).

Results confirmed that both performance and blood flow were sensitive to the experimentally-manipulated task parameters. The key finding was that, although right-hemisphere blood flow was initially higher when searching for feature-presence than for feature-absence, the temporal decline in blood flow during the first 10 minutes of the task was greater in the feature-absence condition. Blood flow remained lower (and continued to decline) in the feature-absence



condition for the remainder of the task. Performance and subjective workload data confirmed that the feature-absence condition was more demanding than feature-presence.

These findings have two principal implications. First, they confirm the applicability of the workload model to feature-absence tasks. Second, they indicate that real-life monitors may have special difficulties in sustaining attention to feature-absence. Blood flow monitoring may be potentially useful in such cases.

*Study 2.* The prototypical vigilance task requires no more than detection of occasional target stimuli, as in the cases of the sentry scanning for enemy combatants, or the radar operator monitoring for enemy planes or submarines. However, many more complex tasks must also be performed for extended periods of time, and may be susceptible to eventual performance decrement. One such task, of military relevance, is vehicle driving. Our previous studies (e.g., Matthews & Desmond, 2002) have shown that the driver's ability to control the vehicle while also monitoring for target stimuli may be impaired following a period of prolonged high workload. The aim for this study was to investigate subjective predictors of performance deficit during simulated vehicle operation, as a precursor to a later study using the TCD. One-hundred-sixty-eight participants performed a simulated drive in our laboratory. Half the sample was exposed to a stress manipulation involving repeated loss of control due to 'black ice'. The study also investigated the role of automation of vehicle speed. Following the stress induction or a nonstressful control drive, participants performed in a dual-task condition requiring the detection of pedestrian stimuli concurrently with driving. Subjective state was assessed using a multidimensional inventory (Matthews et al., 2002) that was administered prior to and following task performance. Results confirmed the efficacy of the stress manipulation, which increased subjective distress. The key finding regarding driver performance was that an index of vehicle control (variability in lane position) was significantly correlated with subjective state factors. Drivers experiencing low task engagement (i.e., tiredness, mental fatigue) and drivers experiencing high levels of worry showed poorer vehicle control.

These findings have two main implications. First, they show that a vehicle driving task is sensitive to a subjective state factor (task engagement) which also predicts vigilance performance (Matthews & Davies, 1998). Sustained attention in the more complex, multi-tasking situation may be controlled by some of the same factors and processes as vigilance. Second, they serve to validate a methodology for investigating predictors of sustained vehicle driving in stressful and non-stressful conditions. We will use these results as the basis for design of a subsequent study investigating the TCD as a predictor of driving performance.

*Study 3* (in progress). Previous TCD studies, including Study 1 of this report, have focused on cerebral blood flow as a concurrent index of functional status, potentially indicating the operator's fitness to perform on a minute-to-minute basis. However, it is important also to *predict* the operator's future performance adequacy. For example, the field commander may need to evaluate whether a soldier is in a fit state to operate equipment or perform a mission requiring sustained attention and vigilance. Study 3 aims to extend our previous work by developing a paradigm for predicting future sustained attention from a short TCD assessment, coupled with additional psychophysiological measures.



In this study, participants perform a battery of three short but highly demanding tasks, requiring signal detection, working memory and psychomotor control, respectively. Such tasks elicit a transient *positive* blood flow response that may index the person's ability to recruit attentional resources and mental effort. To date, little is known about the utility of the response as a predictor of individual behavior. Following this phase of the study, participants perform a longer vigilance task - a 40-minute simulated air traffic control task, used in our previous TCD research (Hitchcock et al., 2003). Prior to performance, we also assess subjective state and salivary cortisol, the latter as a physiological stress index. The design of the study allows us to compare TCD, salivary cortisol, and subjective state as predictors of subsequent vigilance performance.

This study is still ongoing, so key initial findings are summarized briefly. First, manipulation checks confirmed the validity of the methods used. Both the short battery and the longer vigilance task appear to impose high workload and elicit subjective stress. We hope to supplement these findings with data on cortisol response, when they are available. The short task elicited an elevation of cerebral blood flow, as expected and sustained performance of the vigilance task is accompanied by a blood flow decrement. Second, psychometric analyses suggest that individual differences in blood flow may be reliably assessed using our protocol and that phasic task responses may be differentiated from resting baseline. Third, although the study shows that subjective state measures correlate with vigilance, as in previous studies, results so far do not show reliable correlations between vigilance and cerebral blood flow. However, there are indications that the correlation is moderated by the participant's level of involvement in the task. Task-related blood flow responses may indicate higher levels of vigilance in those participants who remain engaged and committed to maintaining task performance.

*Georgia State University.* Owing to delays in local grant administration, hardware installation, and a continuing problem with TCD software, full-scale studies have yet to be initiated. However, extensive pilot work on task development has ensured that data collection using the TCD will begin early in 2005. We have tested over 150 undergraduate participants in pilot research, with data collection and analysis continuing through December, 2004. These studies were designed to validate our shoot/don't-shoot ("Watchkeeper") task and to determine the parameters of testing for other cognitive assessments, such that we are maximally positioned to collect the best data once the TCD is working properly. One study is ongoing. It involves a comparison of response-apparatus options for the Watchkeeper task. Pilot data suggest that the task shows a temporal decrement in performance towards the end of the 20-minute period of work, which may correspond to the conventional vigilance decrement. The other studies were designed to define the optimal Assessment Software for Attention Profiles (ASAP) battery and the relation of this battery to other measures. Pilot data from these studies suggest that measurement of the multiple attentional components specified by Washburn, Putney, and Tirre's (1999) model of attentional abilities has been usefully refined and shortened. Evidence for the validity of the modified ASAP battery has been provided by data suggesting that (1) ability profiles relate to performance on the Watchkeeper task and (2) the ASAP predicts sensitivity of performance to a workload manipulation.

In the remainder of this section of the report, we review completed and ongoing studies in more detail and discuss implications for the timeline of the project.



## STUDY 1. CEREBRAL HEMODYNAMICS AND PERFORMANCE: EFFECTS OF FEATURE PRESENCE/ABSENCE AND WORKLOAD<sup>1</sup>

### *The Problem of Vigilance*

Vigilance tasks require observers to focus their attention and to detect transient and infrequent targets over prolonged periods of time (Davies & Parasuraman, 1982; Warm, 1984; 1993). The first systematic investigation of vigilance dates back to World War II when the Royal Air Force commissioned Norman Mackworth to study an unanticipated and potentially perilous finding in the performance of airborne British Radar Observers on anti-submarine patrol over the Bay of Biscay. After only about 30-minutes on watch, the well-trained and highly motivated observers began to miss "blips" on their pulse-position radar displays that signified surfaced enemy submarines in the sea below. As a result, the undetected U-boats were free to prey upon allied shipping.

Using a simulated task, Mackworth (1950/1961) was able to chart the course of vigilance performance over time and to confirm in the laboratory the suspicions generated in the field that the quality of sustained attention is fragile, waning quickly over time. The progressive decline in vigilance performance noted in Mackworth's original experiments has been termed the "*vigilance decrement*". It has been confirmed repeatedly in subsequent investigations and is the most ubiquitous finding in vigilance research (Davies & Parasuraman, 1982; Warm, 1984). Since Mackworth's pioneering studies, both psychophysical and neurophysiological approaches have been adopted in efforts to understand the nature of sustained attention (Warm, 1984). The present study features a blend of these approaches.

Vigilance also plays a role in automated human-machine systems (Johnson & Proctor, 2004). As Sheridan (1980) has noted, the development of automatic control and computing systems for the acquisition, storage, and processing of information has altered workers' roles from active controllers to a more executive function wherein they must passively attend to a wide variety of displays and actively intervene only in the event of potential problems. Hence, vigilance is a crucial aspect of the reliability of human performance in a wide range of activities such as industrial quality control, air-traffic control, airport security, military surveillance, nuclear power plant regulation, robotic manufacturing, and long-distance driving (e.g., Hancock & Hart, 2002).

One of the reasons for automating these types of assignments is to reduce the information-processing load placed upon operators (Parasuraman & Riley, 1997). However, such reduction appears to be a two-edged sword. Studies have shown that in many instances large-scale accidents can be traced to operators' poor vigilance in the supervision of automated systems. Hence, knowledge of the factors that control vigilance performance is critical for system integrity and public safety (Warm & Dember, 1998).

---

<sup>1</sup> This report is based on T. Hollander's (2003) doctoral dissertation for the University of Cincinnati, available in its entirety on request.



## *Psychophysical Issues*

An important component of research into the nature of sustained attention has been the precise determination of the stimulus characteristics that influence performance. A substantial psychophysical database has emerged demonstrating that the quality of sustained attention depends upon the sensory modality and salience of the critical signals to be detected, as well as upon the frequency, temporal regularity, and spatial regularity of critical signal appearances (Warm, 1993; Warm & Jerison, 1984).

These psychophysical parameters may be important because of their effects on mental workload. Several experiments have reported high levels of perceived mental workload in vigilance tasks, arising from the need to make continuous signal/noise discriminations (Johnson & Proctor, 2004; Warm & Dember, 1998). The vigilance decrement in perceptual sensitivity seems also to be linked to workload; the magnitude of the decrement increases with the difficulty of the signal discrimination (See, Howe, Warm, & Dember, 1995). Prolonged performance of high workload tasks may lead to a loss of attentional resources or capacity, which, in turn leads to vigilance decrement.

The practical application of these findings is that the vulnerability of a task to performance decrement can be assessed by examining task parameters that influence workload. The present study aimed to extend existing findings by examining two further parameters: *search asymmetry* (Quinlan 2003; Treisman & Gormican, 1988) and *event rate*. Search asymmetry refers to the finding that detections are more rapid when searching for the presence of a distinguishing feature in an array of stimuli as opposed to its absence. Indeed, when searching for presence, the distinguishing feature appears to be so salient that it seems to "pop out" of the display. The phenomenon of search asymmetry has been accounted for by the feature integration model (Treisman & Gormican, 1988), which suggests that searching for the presence of a feature is guided by preattentive, parallel processing whereas more deliberate, serial processing is required for determining its absence. Studies by Schoenfeld and Scerbo (1997; 1999) have extended the psychophysics of vigilance by incorporating the presence/absence distinction inherent in the search asymmetry effect into vigilance performance and the perceived mental workload associated with vigilance tasks. As predicted, signal detectability was lower for feature-absent targets than for feature-present targets, when observers were required to search a multi-element display. Subjective workload was also higher for the feature-absent task.

Event rate refers to the overall rate at which stimuli are presented, including both targets and the neutral background events in which they are embedded. The key role of event rate in vigilance has been recognized for many years. A large number of investigations, using both visual and auditory tasks, have demonstrated that the overall level of performance efficiency in vigilance varies inversely with event-rate and that the vigilance decrement is augmented by increments in event rate (Matthews, Davies, Westerman, & Stammers, 2000; Warm & Dember, 1998). This finding is consistent with a workload model of vigilance, in that it may be assumed that fast event rates drain more resource capacity than slow event rates because of the need, at fast event rates, to make more frequent and rapid decisions about whether or not a stimulus event constitutes a critical signal for detection (Davies & Parasuraman, 1982).



## *Brain Systems in Vigilance*

*Brain Imaging via PET and fMRI.* Recent research on vigilance has centered on sophisticated brain-imaging techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) to identify brain systems involved in vigilance. Several systems appear to be activated during task performance, especially those located in the right hemisphere. PET studies have implicated brainstem structures (activation of the midbrain tegmentum and the right laminar region of the thalamus: Kinomura, Larsson, Gulyas, & Roland, 1996), and reduced activation of the cingulate gyrus (Posner & Peterson, 1990). Further brain-imaging studies, reviewed by Parasuraman, Warm, and See (1998) have reported increased blood flow and greater metabolic activity in the right as compared to the left frontal lobe during vigilance task performance. The laterality effects noted in these brain-imaging studies support the results of earlier psychophysical investigations that pointed to right hemispheric specialization in vigilance (Warm et al., 1980).

Although the PET and fMRI studies have been successful in identifying brain regions that are implicated in vigilance, they are subject to a potentially serious limitation: typically they have not related the structures/systems they have identified to performance efficiency as concurrently measured. With the exception of PET studies by Coull, Frackowiak, and Frith (1998) and by Paus and his associates (1997), in which the vigilance decrement was accompanied by PET-measured declines in blood flow in several subcortical structures and in the orbital frontal cortex, the question of the *functional* role of the metabolic changes in brain systems identified in the imaging studies remains largely unanswered (Parasuraman et al., 1998). In addition, PET and fMRI methods may be less than optimal for investigating vigilance because of (1) their high costs, (2) the tendency of subjects to 'fidget' during vigilance (Galinsky, Rosa, Warm, & Dember, 1993), threatening the requirement to stay motionless during brain-imaging, and (3), in the case of fMRI, the loud noise produced by the scanner which may itself make a serious impact on vigilance performance. Therefore, the conditions required for effective use of the PET and fMRI techniques may not provide a suitable environment for correlating changes in brain physiology with vigilance performance over a prolonged period of time.

*The Transcranial Doppler Alternative.* An alternative technique to PET and fMRI for investigating brain systems in vigilance is Transcranial Doppler Sonography (TCD). TCD is a non-invasive neuroimaging technique that employs ultrasound to monitor cerebral blood flow in the mainstem intracranial arteries (Aaslid, 1986). As such, TCD permits inexpensive, continuous, and prolonged monitoring of cerebral blood flow activity during task performance. This technique uses a high-frequency probe (2 MHz) positioned on the scalp to detect blood flow activity in the cerebral arteries. The probe is placed just above the zygomatic arch along the temporal bone, a part of the skull that is transparent to ultrasound, thereby allowing for the identification of blood flow activity in the middle, anterior, and posterior cerebral arteries (Aaslid, 1986). When an area of the brain becomes metabolically active, as in the performance of mental tasks, by-products of this activity, such as carbon dioxide, increase. This results in elevation of blood flow to the region to remove the waste product (Aaslid, 1986; Risberg, 1986).

The main limitation of TCD application is related to its low spatial resolution. In contrast to the regional detail provided by PET and fMRI, TCD as currently employed, can only supply



hemispheric data and cannot provide information about changes in blood flow activity in specific brain loci. However, TCD offers good temporal resolution, permitting tracking of rapid changes in brain blood flow dynamics that can be followed in real time under less restrictive and invasive conditions than PET or fMRI (Aaslid, 1986).

In the past, TCD was used primarily in medicine for neurological diagnosis and for detecting the presence of intracranial vascular dysfunction. Recent studies have indicated, however, that blood flow velocities in the middle cerebral artery (MCA) can be influenced by a variety of tasks such as reading aloud, listening to music, associating words, detecting verbal similarities, detecting and remembering a set of letters, constructing grammatically correct sentences, solving mathematical problems, mentally rotating images, performing a visual search task, and making ethical decisions (Duschek & Shandry, 2003; Klingelhöfer, Sander, & Wittich, 1999; Stroobant & Vingerhoets, 2000; Vingerhoets & Stroobant, 1999). In general, these cognitive activities accelerate blood flow velocity over resting baseline, and the TCD-measured changes are correlated with cognitive demand imposed by those tasks. Because of its good temporal resolution, its ability to measure blood flow in the two cerebral hemispheres, and its sensitivity to task demands, the TCD technique seems to be ideally suited as a tool for meeting the need identified by Parasuraman et al. (1998) to relate brain imaging in vigilance to performance efficiency.

*TCD and Task Type.* An initial study using the TCD procedure (Mayleben, 1998) was guided by a resource-utilization model of vigilance (Warm & Dember, 1998). According to that model, a limited-capacity information-processing system allocates resources or reservoirs of energy to attentionally-demanding processing. The vigilance decrement reflects the depletion of information-processing resources that cannot be replenished in the time available. Mayleben (1998) suggested that cerebral blood flow might represent a metabolic index of resource utilization. Hence, he hypothesized that the vigilance decrement should be accompanied by a decline in cerebral hemovelocity and that the absolute level of blood flow should vary directly with task demand.

To explore these possibilities, Mayleben (1998) employed successive (absolute judgment)- and simultaneous (comparative judgment)-types of psychophysical tasks during a 30-min session. Considerable evidence indicates that the former are more capacity-demanding than the latter, because they make greater demands on working memory (Davies & Parasuraman, 1982). Observers in this experiment were asked to monitor the repetitive presentation of pairs of vertical lines: critical signals for detection in the simultaneous task were cases in which one of the lines was 2mm taller than the other. In the successive task, critical signals were cases in which both lines were 3mm taller than usual. Pilot work insured that the tasks were equated for difficulty under alerted conditions. As expected, the vigilance decrement in this study was accompanied by a parallel decline in cerebral hemovelocity. That decline was more pronounced in the successive than in the simultaneous task. Also consistent with expectations from a workload or resource model, the absolute level of blood flow was significantly higher for observers who performed the more capacity demanding successive task than for those who performed the simultaneous task. In addition, the task-related blood flow results were lateralized – hemovelocity was greater in the right than in the left hemisphere principally in the performance of the memory-based successive task. Mayleben (1998) also employed a control for any effects of boredom, tiredness or repeated



exposure to the stimuli by requiring a further group of observers to watch the dual-line display for 30 min in the absence of a work imperative. Blood flow remained stable over the entire testing period under such conditions.

*Additional TCD studies.* Signal detection in vigilance can be improved by providing observers with cues to the imminent arrival of critical signals (Hitchcock, Dember, Warm, Moroney, & See, 1999). Cued observers need to monitor a display only after having been prompted about the arrival of a critical signal. Therefore, they could husband their information-processing resources over time, whereas non-cued observers were obliged to consume their resources continually. Hitchcock et al. (2003) carried out a TCD experiment in which they varied reliability of a cue that forewarned observers about the imminent arrival of critical signals. Observers monitored a simulated air-traffic control display for planes on a collision course under conditions in which forewarnings about the appearance of critical signals were either 100%, 80%, or 40% reliable, or did not occur at all. Performance efficiency remained stable over time in the 100% reliable condition and declined differentially over time in the other conditions, with the largest vigilance decrement occurring with no cue. As measured by the TCD technique, blood flow in the left cerebral hemisphere declined over time, but it was unrelated to cueing reliability. In the right hemisphere, however, the hemovelocity scores closely mirrored the effects of cueing on performance efficiency. The hemovelocity scores for the several experimental conditions were similar to each other during the early portion of the vigil, but showed differential rates of decline over time, so that by the end of the vigil, blood flow was highest in the 100% group, followed in order by the 80%, 40%, and no-cue groups. As in the Mayleben (1998) study, the observed changes in hemovelocity were dependent on task performance rather than exposure to the display; blood flow remained stable in both cerebral hemispheres when observers were exposed to the simulated air-traffic control display for 40 minutes without a work imperative.

Another recent study using the TCD technique investigated the role of signal regularity, i.e. whether the temporal intervals between critical signals for detection occur in a regular and predictable fashion as opposed to an irregular and unpredictable one. Performance is generally enhanced in the regular signal condition, because observers form expectancies that allow for reduction of workload, although the signal regularity effect may be limited to situations in which signal salience is low (Davies & Tune, 1969). Hollander et al. (2003) asked observers to monitor a simulated air-traffic control display similar to that employed by Hitchcock et al. (2003) for planes on a collision course. Consistent with expectations, perceptual sensitivity was superior for the regular as compared to the irregular critical signal condition but *only* in the context of low salience signals. Blood flow velocities in the left hemisphere, as measured by TCD, showed a general decline in hemovelocity that was unrelated to signal salience or regularity. Blood flow velocities also declined over time in the right hemisphere and were greater in the regular as compared to irregular signal condition. As in the two prior experiments, blood flow remained stable in both cerebral hemispheres among control observers who were exposed to the simulated air-traffic control display for 40 minutes without a work imperative.

Beam et al. (2002) tested whether the relation between vigilance decrement and the right hemisphere decline in blood flow generalized to a cognitive vigilance task, i.e., a task requiring a symbolic target discrimination rather than a sensory discrimination. Observers performed either a sensory task, a simple cognitive task, or a complex cognitive task. Blood flow in the right middle



cerebral artery paralleled performance, indexed in terms of the accuracy of signal detections, for all three tasks. The sensory task showed both a significant decrement in performance and a concurrent decline in blood flow. The two cognitive tasks showed stable performance levels, and neither task showed a significant decline in right-hemisphere blood flow. Given that cognitive tasks sometimes show significant vigilance decrement (See et al., 1995), this results points towards the need to further investigate cognitive tasks, a goal for a later study in the present research.

### *Aims and Hypotheses*

The recent studies we have reviewed demonstrate that decline in right-hemisphere blood flow provides a reliable index of loss of performance capability on several different tasks requiring sustained attention. Conversely, blood flow in both hemispheres remains stable when performance does not show temporal decline. The present study aimed to further generalize these findings to the case in which the target is defined by the absence of a specified stimulus element. Specific goals for the study were as follows.

Given that a fast event rate is more capacity-demanding than a slow event rate and that detecting the absence of a feature is more capacity-demanding than detecting its presence, the first goal for the current study was to test the expectation that vigilance decrement would be most evident when the target was defined by the absence than the presence of a critical stimulus feature, and event rate was high. As in the studies by Schoenfeld and Scerbo (1997; 1999), the same outcome was anticipated with regard to perceived mental workload. Consequently, the second goal for this study was to test that expectation. The remaining goals pertained to the validity of blood flow as an index of compromised attention. Given that blood flow in vigilance and a variety of other tasks has been found to vary directly with the information-processing demand imposed upon the observer (Hitchcock et al., 2003) and that detecting the absence of a stimulus feature is more capacity-demanding than detecting its presence, it was anticipated that blood flow would be higher in the absence than the presence case. Moreover, given the findings by Hitchcock et al. (2003) and Mayleben (1998) that the temporal decline in blood flow was directly related to task demand, it was also anticipated that blood flow should show a greater decline over time in the absence than in the presence condition. Finally, based on prior findings of right hemisphere dominance in vigilance, it was also anticipated that the blood flow effects associated with the presence/absence distinction would be restricted to the right hemisphere. Testing these expectations were the third, fourth, and fifth goals for this study.

## **Method**

### *Participants*

One-hundred-twenty-eight undergraduate students (64 men and 64 women) from introductory psychology classes served as observers for course credit. They ranged in age from 19-30 years, with a mean age of 22.4 years. All had normal or corrected-to-normal vision and were right-handed, as indexed by a greater proportion of right- than left-preference responses on the Edinburgh Handedness Inventory (Oldfield, 1971).



## Design

Sixteen observers (8 men and 8 women) were assigned at random to each of six active experimental conditions defined by the factorial combination of two levels of task type (presence, absence) and three levels of event rate (6, 12, 24 events/min). Due to the unilateral imaging capability of the TCD equipment, half of the participants in each of the active experimental groups, equated for sex, had TCD recordings taken from the left cerebral hemisphere, while the remainder had recordings taken from the right cerebral hemisphere.

The remaining 32 observers served as passive controls to ensure that time-based changes in blood flow were indeed task-linked and not artifacts of merely sitting in the experimental room for an extended period of time. These observers simply looked at the displays (see below) for 40 min without an information-processing imperative. Two groups of 16 passive observers (equated for sex) were selected at random to have hemovelocity measures taken from the left or right MCAs. In order to provide an equal number of observers in all control conditions, the two extreme event-presentation conditions were employed in the control groups. Within each control group, 4 observers, equated for sex, were exposed to the presence display format and the remainder to the absence format as described below.

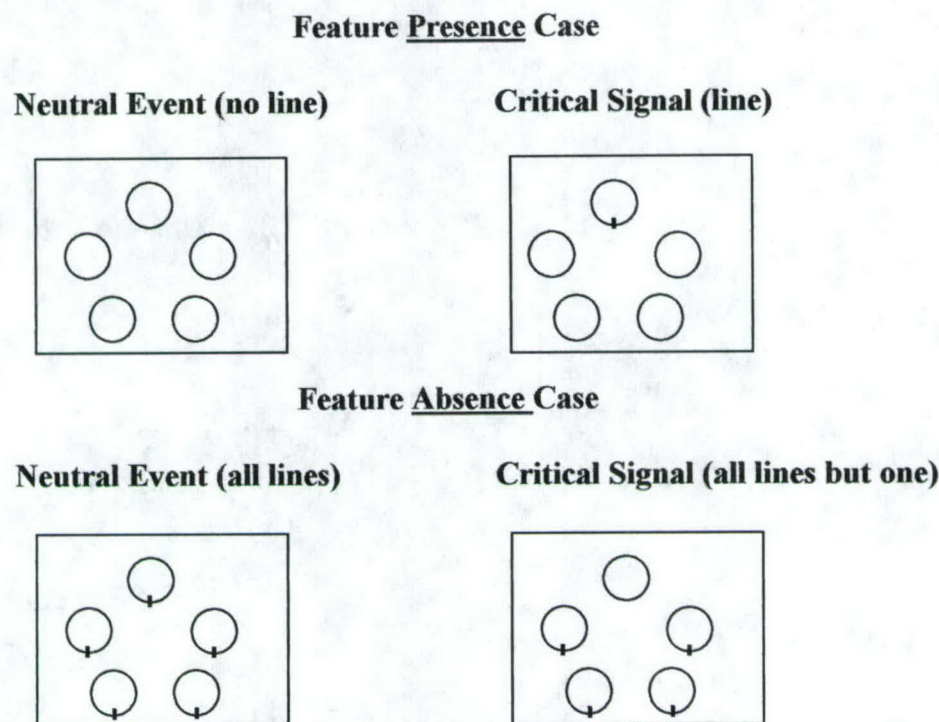
## Apparatus

In all experimental conditions, observers participated in a continuous 40-min vigil divided into four 10-min periods of watch during which they monitored an array of five circles (14 mm diameter) positioned around the center of a video display terminal (VDT) at the 3, 5, 7, 9, and 12 o'clock locations. Examples of the feature presence and absence conditions are illustrated in Figure 1.

The critical signal for detection in the presence condition was the appearance of a vertical 4 mm line intersecting the 6 o'clock position within *one* of the circles in the array. In the absence condition, the vertical 6 o'clock line was present in all circles *but one*. Ten critical signals were presented in each watchkeeping period in all experimental conditions. The schedule of critical signal appearances was varied at random for each observer in each condition, with the restrictions that signals appeared on an average of once/min during each period of watch and that the signals appeared equally often in each of the five circles comprising the vigilance display during each period of watch. All stimuli were exposed for 0.40 sec. Event rates of 6, 12, and 24 events/min were achieved by setting stimulus onset asynchronies (SOAs) at 10 sec, 5 sec and 2.5 sec, respectively. In all experimental conditions observers signified their detection of critical signals by pressing the spacebar on a computer keyboard. Responses occurring within 1.5 sec after the appearance of critical signals were recorded as correct detections; all other responses were classified as false alarms. Pilot work ensured that signals in the presence and absence conditions were equally detectable under alerted conditions.

Perceived mental workload was measured by a computerized version of the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988), which was administered immediately after the vigilance session. This instrument is one of the most effective measures of perceived mental





Note: Critical signals can appear in any of the five circles.

Figure 1. Examples of neutral and critical events in feature presence/absence conditions.

workload currently available (Wickens & Hollands, 2000). It was the workload index employed by Schoenfeld and Scerbo (1997; 1999) in their introduction of the feature presence/absence distinction to vigilance research.

Hemovelocity measurements, percent blood flow relative to a baseline index described below, were taken from either the right or left middle cerebral arteries by means of a DWL/Multi-Dop X4 TCD unit, as illustrated in Figure 2. Observers wore a 2-MHz ultrasound transducer embedded in a plastic bracket which was secured around the head by an adjustable Velcro strap and located dorsal and immediately proximal to the zygomatic arch along the temporal bone. A small amount of Aquasonic-100 brand ultrasound transmission gel (Parker Laboratories, NJ, USA) was placed between the TCD transducer and the observer's skin to enhance the blood flow signal. The mainstem MCA, which was monitored at depths of 45-55mm (distance between the transducer face and sample volume), typically has a hemovelocity range of 50-65 cm/sec. The TCD unit permitted depth adjustment in 5mm increments as needed for isolating the MCA. Time-averaged blood flow velocities were displayed by the unit every four seconds. These values were channeled into a PC for subsequent data analysis.

Baseline hemovelocity measures were obtained from the appropriate MCA (left or right) during a 5-minute period in which observers were seated in front of the blank VDT. All active observers wore the transducer throughout the subsequent practice and testing phases; however, recordings were secured only during the testing phase. In the case of the control observers, baseline measurements were followed immediately by the 40-min passive observing session.





Figure 2. Illustration of the transcranial Doppler unit (TCD).

### *Procedure*

Observers were tested individually in a windowless laboratory room. The VDT was located at eye-level on a table sited 60 cm directly in front of the seated observer. Upon reporting for the experiment, all observers received a verbal briefing from the experimenter and then completed an informed-consent form and the Edinburgh Handedness Inventory. Afterwards, they were acclimated to the TCD procedure, and observers in the active experimental groups were given two 5-min practice trials, followed by the main, 40-minute vigilance task.

## **Results**

### *Vigilance Performance*

*Detection Probability.* Mean percentages of correct detections in all experimental conditions are presented in Table 1. Perusal of the table will reveal that detection probability was greater in the feature presence ( $M = 75.8\%$ ,  $SE = 1.55$ ) than in the feature absence condition ( $M = 73.4\%$ ,  $SE = 1.98$ ) and that detection probability generally declined over time. These impressions were confirmed by a 2 (task type)  $\times$  2 (event rate)  $\times$  4 (periods of watch) analysis of variance (ANOVA) performed on the arcsines of the percentage scores. The arcsine transform was employed to normalize the percentage data (Kirk, 1995). The ANOVA revealed significant main effects for task type,  $F(1, 90) = 16.27$ ,  $p < .01$ , and periods of watch,  $F(2, 90) = 4.89$ ,  $p < .01$ . In addition, there was a significant Event Rate  $\times$  Task Type interaction,  $F(2, 90) = 4.52$ ,  $p < .02$ . All other sources of variance in this analysis lacked significance,  $p > .05$  in each case. In this and all subsequent ANOVAs, Box's epsilon was employed when appropriate to correct for violations of the sphericity assumption (Maxwell & Delaney, 2004). The overall mean false alarm rate was only 3.5%, and, consequently, the false alarm data were not analyzed further.



Table 1. Mean percentages of correct detections for the two task types at each event rate within each period of watch. Standard errors are in parenthesis.

Task Type	Event Rate (Events/Min)	Periods ( 10 minutes)				M
		1	2	3	4	
Presence	6	77.0 (1.23)	75.4 (1.42)	75.4 (1.72)	74.8 (1.85)	75.7 (1.56)
	12	78.1 (1.54)	75.9 (1.78)	75.1 (1.13)	75.0 (1.64)	76.1 (1.52)
	24	76.7 (1.37)	75.9 (1.64)	74.9 (1.49)	74.7 (1.77)	75.6 (1.57)
Absence	6	78.5 (1.47)	75.2 (1.86)	75.1 (2.14)	74.3 (2.54)	75.8 (2.01)
	12	75.7 (1.37)	72.7 (1.65)	71.1 (1.47)	69.6 (2.78)	72.3 (1.81)
	24	72.9 (1.54)	73.1 (1.92)	72.2 (2.42)	69.9 (2.67)	72.1 (2.13)
M		76.5(1.42)	74.7 (1.71)	73.9 (1.72)	73.1 (2.20)	

The Event Rate  $\times$  Task Type Interaction is illustrated in Figure 3, wherein detection probabilities are plotted as a function of event rate for each task-type condition. It is evident in the figure that performance efficiency declined with increments in event rate in the absence condition, while event rate had little effect upon signal detection in the presence condition. Separate ANOVAs on the data of the two task conditions revealed a significant effect for event rate in the absence condition,  $F(2, 45) = 9.37$ ,  $p < .01$ , but not in the presence condition,  $F(2, 45) < 1$ .

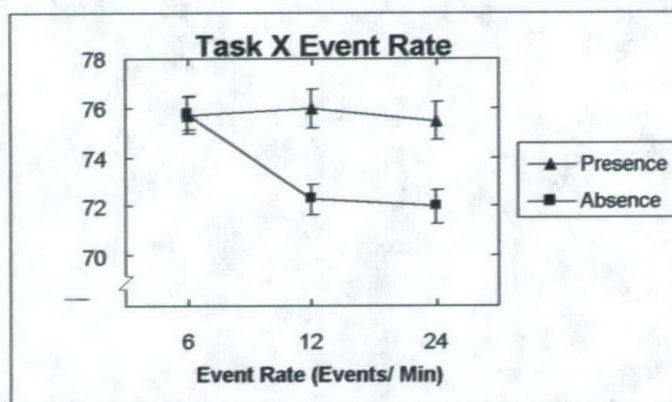


Figure 3. Mean percentages of correct detections as a function of event rate for the feature presence/absence conditions. Error bars are standard errors.

### NASA Task-Load Index

Workload scores for the active observers fell at the mid to upper level of the NASA-TLX scale and were considerably higher than those for the passive control observers, whose mean was 37.3 (3.64), as shown in Table 2. Among the active observers, the workload was higher for those who monitored for feature absence than for feature presence. A 2 (task type)  $\times$  3 (event rate) ANOVA of the global workload scores revealed a significant main effect for task type,  $F(1, 90) = 5.35$ ,  $p < .05$ . The main effect for event rate and the Task Type  $\times$  Event Rate interaction were not significant,  $p > .05$  in each case.



Table 2. Mean global workload scores for the feature-presence and feature-absence groups at each event rate and for the passive-control group. Standard errors are in parenthesis.

Condition	Event Rate (Events/Min)			M
	6	12	24	
Presence	51.3 (4.51)	48.2 (4.83)	57.1 (5.21)	52.2 (4.85)
Absence	62.7 (3.92)	60.3 (4.21)	58.7 (4.64)	60.6 (4.26)
M	57.0 (4.21)	54.3 (4.52)	57.9 (4.93)	

### Cerebral Hemovelocity

**Control Issues.** To take account of the wide range of hemovelocity scores present in the population (Adams, Nichols, & Hess, 1992), the hemovelocity scores for all observers in this study were expressed as a proportion of the last 60-sec of their 5 min resting baseline. Baseline scores were similar in all experimental groups and in the passive control groups. In addition, a 2 (hemisphere)  $\times$  4 (periods of watch) mixed analysis of variance (ANOVA) of the control observers' hemovelocity scores showed no effects for time on task in either hemisphere,  $p > .05$  for all sources of variance. Thus, time-related blood flow effects in the active performance groups cannot be attributed to factors associated with remaining seated in the experimental room for a prolonged period of time. Control observers' mean hemovelocity scores in the right and left cerebral hemispheres are plotted as a function of time on task in Figure 4.

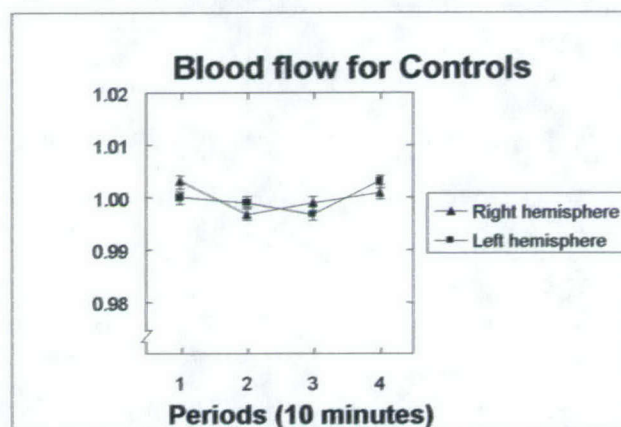


Figure 4. Control observers' mean hemovelocity scores in the right and left cerebral hemispheres as a function of time on task. Error bars are standard errors.

**Hemovelocity During Active Vigilance.** Mean hemovelocity scores in all experimental conditions are presented in Table 3. A 2 (task)  $\times$  2 (hemisphere)  $\times$  2 (event rate)  $\times$  4 (periods of watch) mixed ANOVA of the blood flow data revealed that the overall level of blood flow was significantly greater in the right ( $M = .98$ ;  $SE = .006$ ) than in the left hemisphere ( $M = .94$ ;  $SE = .006$ ),  $F(1, 84) = 20.99$ ,  $p < .01$ ; and significantly higher in the feature *presence* condition ( $M = 0.97$ ;  $SE = .006$ ) than in the feature *absence* condition ( $M = 0.95$ ;  $SE = .006$ ),  $F(1, 84) = 8.23$ ,  $p < .01$ ; and that the overall level of blood flow declined significantly over time. ( $M$ 's for periods 1 through 4 = 1.00, 0.96, 0.95, and 0.92, respectively,  $SE$ 's = .021, .019, .019, and .021),  $F(3,$



252) = 56.01,  $p < .01$ . In addition, there were two significant interactions: Task  $\times$  Hemisphere,  $F(1, 84) = 4.42$ ,  $p < .05$  and Periods  $\times$  Hemisphere,  $F(3, 252) = 5.70$ ,  $p < .01$ . All other sources of variance in the analysis lacked statistical significance,  $p > .05$  in each case.

Table 3. Mean hemovelocity in the left and right MCA's for all combinations of task type and event rate during each period of watch. Standard errors are in parenthesis.

Hemisphere	Task Type	Event Rate (Events/Min)	Periods (10 minutes)				M
			1	2	3	4	
Left	Presence	6	<b>1.02</b> (.021)	<b>0.94</b> (.019)	<b>0.91</b> (.019)	<b>0.90</b> (.021)	<b>0.94</b> (.020)
			<b>1.01</b> (.024)	<b>0.95</b> (.012)	<b>0.94</b> (.023)	<b>0.90</b> (.026)	<b>0.95</b> (.021)
		12	<b>0.97</b> (.021)	<b>0.93</b> (.019)	<b>0.93</b> (.019)	<b>0.90</b> (.021)	<b>0.93</b> (.020)
		24					
	Absence	6	<b>0.99</b> (.021)	<b>0.94</b> (.019)	<b>0.91</b> (.024)	<b>0.88</b> (.021)	<b>0.93</b> (.022)
			<b>1.00</b> (.023)	<b>0.95</b> (.019)	<b>0.94</b> (.019)	<b>0.91</b> (.025)	<b>0.95</b> (.020)
		12	<b>0.97</b> (.021)	<b>0.94</b> (.019)	<b>0.91</b> (.021)	<b>0.87</b> (.021)	<b>0.92</b> (.021)
		24					
	Presence	6	<b>1.02</b> (.024)	<b>1.01</b> (.019)	<b>1.00</b> (.021)	<b>0.99</b> (.021)	<b>1.01</b> (.021)
			<b>1.02</b> (.021)	<b>0.99</b> (.019)	<b>0.99</b> (.019)	<b>0.97</b> (.021)	<b>0.99</b> (.020)
		12	<b>1.02</b> (.024)	<b>1.00</b> (.019)	<b>1.00</b> (.019)	<b>0.99</b> (.021)	<b>1.00</b> (.021)
		24					
	Absence	6	<b>1.00</b> (.021)	<b>0.99</b> (.023)	<b>0.98</b> (.019)	<b>0.95</b> (.027)	<b>0.98</b> (.021)
			<b>0.99</b> (.024)	<b>0.96</b> (.019)	<b>0.91</b> (.019)	<b>0.87</b> (.021)	<b>0.93</b> (.022)
		12	<b>0.98</b> (.021)	<b>0.96</b> (.019)	<b>0.96</b> (.012)	<b>0.94</b> (.024)	<b>0.96</b> (.020)
		24					
		M	<b>1.00</b> (.021)	<b>0.96</b> (.019)	<b>0.95</b> (.019)	<b>0.92</b> (.021)	

The Task  $\times$  Hemisphere interaction is presented in Figure 5. Mean blood flow scores for the target presence/absence conditions are displayed for each hemisphere. Separate ANOVAs of the scores for the two discrimination conditions revealed that blood flow in the right hemisphere was significantly higher when the critical signals for detection were defined by feature presence than absence,  $F(1, 42) = 11.21$ ,  $p < .01$ . Presence/absence differences were negligible in the left hemisphere,  $F(1, 42) < 1$ .



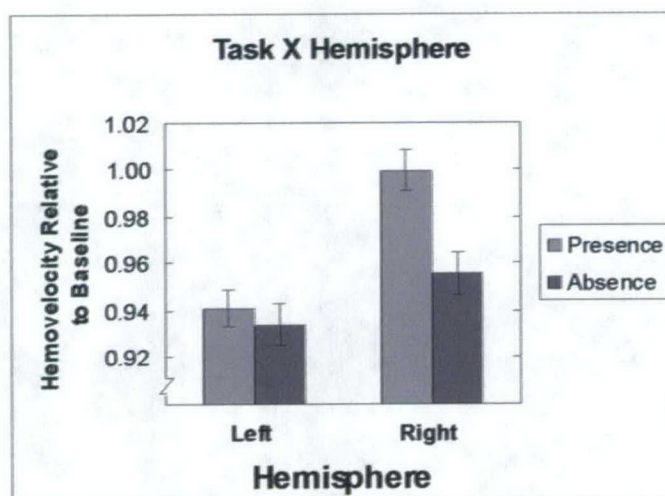


Figure 5. Mean hemovelocity scores in the right and left cerebral hemispheres for the feature presence/absence conditions. Error bars are standard errors.

The Periods  $\times$  Hemisphere interaction is shown in Figure 6. Blood flow values for each hemisphere are plotted as a function of periods of watch. While blood flow in each hemisphere declined significantly over time,  $F_{\text{right hemisphere}}(2, 103) = 13.58, p < .01$ ;  $F_{\text{left hemisphere}}(2, 95) = 47.82, p < .0$ , it is evident in the figure that the rate of decline was greater in the left than in the right hemisphere.

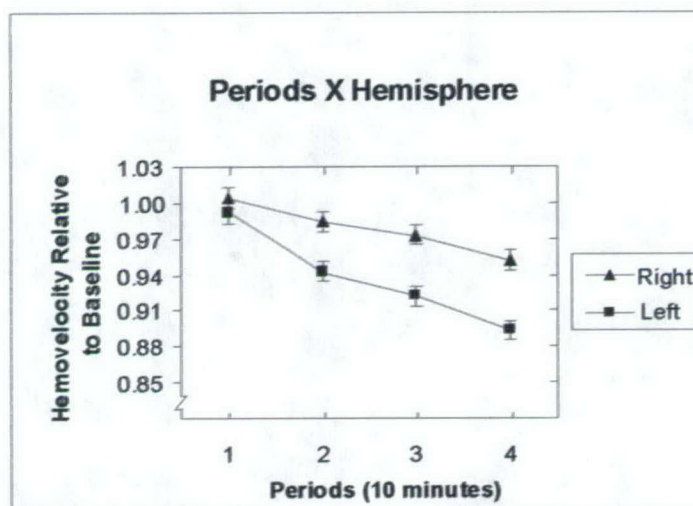


Figure 6. Mean hemovelocity scores in the right and left cerebral hemispheres as a function of time on task. Error bars are standard errors.

The finding that blood flow was greater in the presence as compared to the absence condition was counter to expectations based on the view that detecting feature absence is more capacity-demanding than detecting feature presence. It is conceivable, however, that this apparent reversal of the expected effect reflects the fact that the demands of feature absence are great enough to tax information-processing resources early in the vigil and that those resources are not replenished over time. An account along this line would be supported if it could be shown that



while differences in blood flow are greater in the feature absent than the feature present case at the outset of the vigil, the reverse effect emerges as the vigil continues.

Toward that end, a fine-grained minute by minute analysis was performed on the blood flow scores of the present and absent conditions during the initial watchkeeping period in the left and right hemispheres. No task differences were noted for the left hemisphere. As shown in Figure 7, however, blood flow in the right hemisphere appears to be greater in the feature *absent* than the feature *present* case at the outset of the vigil and the reverse effect emerged after observers performed the task for six minutes.

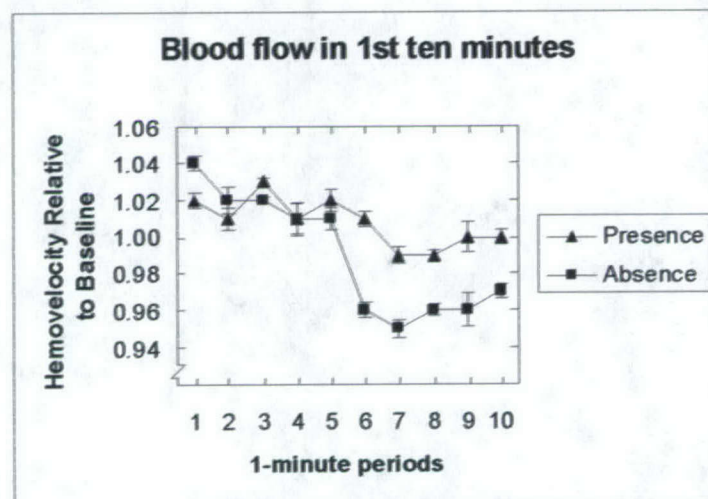


Figure 7. Mean hemovelocity scores in the feature presence/absence conditions for successive one-minute intervals during the initial period of watch. Data are for the right cerebral hemisphere. Error bars are standard errors.

An ANOVA of the data for the right hemisphere revealed significant main effects for periods,  $F(7, 280) = 23.08, p < .01$ , and discrimination condition,  $F(1, 42) = 27.07, p < .01$ , and a significant interaction between these factors,  $F(7, 280) = 8.90, p < .01$ . Supplementary *t*-tests of the differences between the absence and presence conditions using an alpha level of .05 and the Bonferroni correction revealed that there were no statistically significant differences between the conditions from the 1<sup>st</sup> through the 5<sup>th</sup> minute of watch. However, there were statistically significant differences between the conditions in the 6<sup>th</sup> through the 10<sup>th</sup> minute of watch.

Similar fine-grained examination of the data for the remaining periods of watch revealed that the reduced level of blood flow in the absence condition that emerged half way through the initial watchkeeping period consistently remained throughout each minute of all of the remaining periods of watch.

### Discussion

**Vigilance Performance and Workload.** One goal for the present study was to test the prediction that the event-rate effect in vigilance would be more pronounced when critical signals for detection were defined by feature absence than by presence. That prediction was based upon the feature-integration model, in which detecting feature absence is considered to be more capacity-demanding than detecting feature presence (Triesman & Gormican, 1988) and upon prior



findings that the event-rate effect is more pronounced in tasks requiring high information-processing demand than low (Warm & Dember, 1998). Consistent with that prediction, detection probability in this study varied inversely with event rate in the absence condition while event rate had no effect upon performance efficiency in the presence condition. This result is reminiscent of the findings in the earlier reports by Schoenfeld and Scerbo (1997; 1999) that the effect of another information-processing factor in vigilance, the size of the element set that must be scanned in search of critical signals, is also more notable in the feature absence than the feature presence condition. Also consistent with Schoenfeld and Scerbo's earlier reports (1997; 1999), perceived mental workload was greater when observers monitored for feature absence than presence. Taken together, the present study and the studies by Schoenfeld and Scerbo support the notion of differential capacity demand in detecting feature absence than presence.

However, the finding of a vigilance decrement in the feature presence condition suggests that, counter to the early claim in feature-integration theory that feature detection is preattentive, some information processing cost is associated with detecting feature presence. This interpretation is supported by the fact that even though the workload of the presence condition was less than that of the absence condition, it still fell at the upper level of the TLX scale. This result is consistent with the emerging view in the search literature that the alignment of feature detection with preattentive processing may no longer be tenable (Quinlan, 2003).

*Cerebral Hemovelocity.* In an effort to provide additional converging evidence that detecting feature absence is more capacity-demanding than detecting feature presence, this study also tested the predictions that cerebral blood flow would be greater in the absence than in the presence condition and that there would be a steeper decline in blood flow over time in the absence than in the presence condition. This expectation was not confirmed when blood flow was measured within 10-min time blocks. However, a more fine-grained minute-by-minute analysis of blood flow indicated that while hemovelocity was similar in the presence and absence conditions at the outset of the vigil, blood flow in the absence condition dropped significantly below that in the presence condition by the sixth minute of watch and continued to fall below that of the presence condition for the remainder of the vigil. Evidently, the initial expectation of greater blood flow in the absence condition underestimated the degree to which that condition taxed information-processing resources in the vigilance task. Rather than being reflected in an overall elevation of blood flow, the greater information-processing demand exerted by the absence condition was evident in an early-appearing drain on resources. Although initially unanticipated, this early-appearing drain in the fine-grained analysis of the data is consistent with the expectation of a steeper decline in blood flow associated with the absence condition.

It is important to note that the blood flow effects associated with the increment in the information-processing load imposed by the feature absence condition were restricted to the right cerebral hemisphere. That result is concordant with previous findings with cerebral blood flow, as well as with PET and fMRI findings, pointing to the operation of a right-hemispheric system in the functional control of vigilance performance (Hitchcock, et al., 2003; Parasuraman et al., 1998). The lateralization of the hemovelocity findings regarding feature presence/absence also has another important implication—it makes it unlikely that the blood flow changes associated with that stimulus dimension were the result of gross changes in systemic vascular activity linked to vigilance performance, such as variations in blood pressure, cardiac output, and changes in heart



rate (Caplan et al., 1990). Such gross changes in vascular activity are not likely to be hemispheric-dependent.

Unlike the blood flow effects associated with the feature presence/absence dimension, the vigilance decrement was accompanied by a decline in blood flow in both hemispheres, more so in the left than in the right. Bilateral declines in blood flow over time were also observed in previous vigilance studies (Beam et al., 2002; Hitchcock et al., 2003; Hollander et al., 2003). Evidently, vigilance performance is not completely lateralized, since both hemispheres appear to have a part in the temporally-linked aspects of vigilance performance, consistent with Hellige's (1993) proposal that even relatively simple tasks require the coordination of several information-processing subsystems.

An issue to consider in regard to cerebral hemovelocity is the absence of any association between blood flow and variations in event rate. The absence of any such relation in the present study may be due to the restricted range of event rates employed, with a maximum rate of 24 events/min. Future research with discriminations that permit a broader range in the pace of stimulus presentation is clearly needed to further explore the potential relation between event rate and blood flow in vigilance.

*Practical Implications.* As described initially, vigilance research was born in the context of an applied problem and continues to have relevance in a wide variety of automated human-machine systems. Thus, it is important to note that in addition to their theoretical meaning, the results of this study and those of the earlier experiments by Schoenfeld and Scerbo (1997; 1999) may have important implications for operational situations in which observers must monitor displays for changes in system function. For example, air-traffic controllers must be on the lookout for the unanticipated disappearance of an icon that can signify a plane in trouble. As Schoenfeld and Scerbo (1997; 1999) have noted, the greater information-processing load imposed by searching for feature absence than presence implies that feature presence/absence is not a matter of indifference where display design is concerned. Whenever possible, displays should require that observers monitor for the presence of important features not for their absence.

In the military context, it is not unusual for monitors to search for the absence of stimuli. Effective surveillance of any enemy grouping such as soldiers, tanks, or ships typically requires the observer to detect the disappearance of any individual unit. Similarly, one potential component of the Land Warrior and Future Force Warrior programs is the delivery to the soldier, via a head-up display, of tactical maps indicating locations of friendly and enemy combatants. Vigilance for the disappearance of personnel from the map would seem critical. The present study highlights the difficulty of sustaining attention to missing stimulus elements, even over the relatively short time period of 40 minutes. It also confirms that monitoring TCD may provide a useful index of loss of functional capacity to perform this task.



## STUDY 2: THE EFFECTS OF STRESS AND AUTOMATION ON PERFORMANCE IN A SIMULATED WINTER DRIVE<sup>2</sup>

### *The Influence of Driver Stress on Performance*

Everyone who drives has experienced stress while driving at some time or another. Hazardous road conditions, dangerous maneuvers of other drivers, and frustration from traffic congestion may all contribute to feelings of stress while driving. Diary studies of driving indicate that drivers frequently experience stress symptoms associated with driving, even during familiar drives such as commuting (Gulian, Glendon, Matthews, Davies, & Debney, 1990). Such symptoms may include unpleasant emotion, worry, and minor health problems (Gulian et al., 1990), together with stronger responses to especially threatening events (Matthews, 2001). Therefore, an understanding of the influence of stress on accident involvement is vital for promoting motor vehicle safety.

In the military context, motor vehicle accidents are a major source of loss of life. FY03 data from the United States Army Safety Center (*Enhancing combat readiness*, n.d.) show that, of 255 army military fatalities, 60% were attributable to vehicle accidents (including those involving privately owned vehicles). Most accidents related to warfighting occurred in convoys in forward areas. A variety of factors contribute to accidents, including fatigue, excessive speed and environmental conditions, but it is not unlikely that stress plays some part. In Iraq, for example, convoys are commonly targeted by insurgents, and, even when a convoy is not under attack, knowledge of this threat may be a source of distraction from driving.

Traditional theories of stress tend to attribute stress solely to properties of the stimulus or properties of the individual's response (e.g., Gray, 1987). However, contemporary theories of stress tend to reject these positions as too simplistic. An alternative to these theories is the transactional model of stress, which defines stress as "a relationship between the person and the environment that is appraised by the person as taxing or exceeding his or her resources and endangering his or her well-being" (Lazarus & Folkman, 1984). According to the transactional model, the experience of stress is not exclusively dependent upon environmental or intrapersonal factors. Instead, stress develops as an interaction between external stressors and the individual's cognitive and behavioral responses to those stressors. The person must first evaluate the demands imposed by external stressors in the light of available coping strategies (Matthews, 2001); stress symptoms such as anxiety develop when demands are appraised as exceeding coping abilities. Thus, response to demanding driving environments reflects not just external stress factors but also the individual driver's personality.

### *The Influence of Automation on Performance*

Another factor that may influence driver performance is vehicle automation. In-vehicle technology is becoming increasingly complex as new devices and systems are added to modern automobiles. By 2015, at least five major automobile manufacturers plan on incorporating systems

---

<sup>2</sup> This report is based on G.J. Funke's (2003) Master's Thesis for the University of Cincinnati, available in its entirety on request.



designed to monitor headway error, track and avoid collisions, and adaptively maintain a set cruise speed (Walker, Stanton, & Young, 2001). Increasing automation of functions previously performed by human drivers is also integral to Future Combat Systems such as armed robotic vehicles, recovery and maintenance vehicles, intelligent munitions systems and unmanned air vehicles (Mouloua, Gilson, & Hancock, 2003).

Parasuraman and Riley (1997) have operationally defined automation as "the execution by a machine agent (usually a computer) of a function that was previously carried out by a human." While there are many reasons for automating a task (see Wickens & Hollands, 2000 for a detailed discussion of these reasons), of particular interest is the claim that automation may reduce operator workload. However, whether or not a task is automated is often determined by technological and economic feasibility, rather than by careful analysis of the consequences (Parasuraman & Riley, 1997). This approach tends to define the operator's role in terms of the automation, often relegating them to positions of supervising and monitoring the automated system, a change which may engender unintended associated consequences. First, automation may not actually reduce operator workload, but instead transfer it to new tasks (Reinartz & Gruppe, 1993). For example, some pilots feel that aviation automation has actually increased their workload, particularly during high-workload phases of the flight (i.e., during the descent and final approach; Wiener, 1989). Second, operators may place too much trust in the automated system (May, Molloy, & Parasuraman, 1993). As a system becomes more reliable, the operator becomes complacent, and spends less time and effort on monitoring the automated system for errors, resulting in failures to detect automation malfunctions.

These problems may be further compounded by the operator's potential loss of situational awareness (Wickens & Hollands, 2000). Situational awareness refers to the operator's mental representation of the automated task, and it encompasses factors that are both internal and external to the task. In the driving context, internal factors include the state of the automated system, the route selected by the driver, whether or not the driver is fatigued or stressed, etc. External factors include local traffic patterns, geographical location, current road conditions, etc. A driver that trusts a reliable automated system may perceive little need to monitor the system, and thereby lose situational awareness of the driving context. In this case, even if the driver does detect an error in the automated system, he or she may not be prepared to deal with the problem appropriately. This loss of situational awareness may impair the driver's judgments about the driving situation, leading them to make critical errors during the driving task.

Currently, the interaction of driver stress and vehicle automation has not been investigated. In general, studies reveal that automation may indeed result in stress for operators, and that prolonged exposure to such stressful conditions may lead to an increased risk of contracting stress-related diseases (e.g., myocardial infarctions, musculoskeletal disorders, etc.) (Lundberg & Johansson, 2000). However, such studies often fail to relate stress responses to task performance indices. In one study that did investigate performance - crew errors during aircraft operation - it was shown that automation as a stressor does have some impact on performance, but that crews were sufficiently adaptive to overcome such difficulties with only minor performance impairments (Kanki, 1996).

However, studies such as the above have not considered the potential joint effects of



stress and automation on performance. Therefore, it is still unclear what the role and impact of stress on performance will be in the presence of automated technologies. The negative impact of driver stress may be mitigated or exacerbated by automation. Two theories exist within the performance literature that may allow us to make some predictions about the effects of stress and automation on performance.

### *Theories of Performance*

*Attentional Resource Theory.* Attentional resource theories (e.g., Norman & Bobrow, 1975) posit that information processing and task performance are dependent on the availability of system resources. Typically these resources are believed to exist in a fixed quantity, though there is some evidence to suggest that this quantity may vary with task demands and other factors (see Young & Stanton, 2002 for a review). These resources act as an energizer for information processing; workload may represent the proportion of resources required to meet the demands of a task (e.g., Wickens & Hollands, 2000). According to resource theory, as task demands increase, more effort is required to perform a task, and therefore, more resources are used up in the performance of the task and workload increases. If the available resources are not sufficient to meet the demands of the current situation, skilled operators may adjust their strategy to compensate, or barring this, performance deficits will be evidenced.

Stressor effects may be mediated by shortfalls in resources, resulting either from a shrinkage of the resource pool, or because some resources must be allocated to processing stimuli associated with the stressor. In support of this viewpoint, various stressors, including noise, heat, anxiety, subjective tiredness, and prolonged work have been shown to impair performance most reliably when a task is attentionally demanding (see Matthews et al., 2000 for a review). Furthermore, stressor effects generalize over an assortment of qualitatively different tasks, indicating that they are affecting some general resource, rather than a single critical process (Matthews & Davies, 1998).

A variety of driving studies have shown a negative relation between task workload and driver performance, supporting a resource model for driver performance. For example, Harms (1986) found that driver workload ratings on 100-meter segments of roadway covaried with the number of reported traffic accidents on those segments. In a partial review of the existing literature on driver workload, Kantowitz and Simsek (2001) report that performance of a secondary task while driving may influence several objective indices of driver performance. Driver workload may also be influenced by demanding roadway geometries such as heavy road curvature and narrow lanes, resulting in impaired driver performance under those conditions.

Thus, stress and automation may have opposing effects on driver performance, stress through reducing the quantity of available resources and hence increasing the driver's vulnerability to high workload, and automation through reducing workload and hence stress vulnerability. According to the resource model, conditions that impose high levels of information processing load should show the greatest performance impairments.

*Effort Regulation Hypothesis.* By contrast with resource theory, the effort regulation hypothesis predicts that stress and automation may operate synergistically, to the extent that both



tend to reduce the driver's task-directed effort. The effort regulation hypothesis attributes performance impairments to a failure to match effort to environmental task demands (Hancock & Warm, 1989). Such performance impairment may be due to a lowering of performance standards (Hockey, 1997), or it may be a consequence of insufficient performance monitoring and lack of awareness of impairment (Matthews & Desmond, 2002). In driving, when the perceived workload of a task is low, drivers may under-estimate the difficulty of current driving conditions, leading them to apply less effort to the task. If the driver is also subject to stress, he or she may also divert attention to personal worries. However, when the perceived workload of the task is high, participants must mobilize additional effort to maintain safety and performance, and this additional effort may focus attention on the task, and mitigate the effects of stress on performance. Previous driver simulator studies have found that stress (Matthews, 1996) and fatigue (Matthews & Desmond, 2002) inductions have more serious adverse effects on vehicle control and attention to hazards during straight sections of roadway, as opposed to curved sections of roadway. These findings support an effort regulation model of stress and performance; contrary to resource theory, stress and fatigue may be most damaging when workload is low.

Therefore, it is possible that stress and automation may have an interactive impact on driver performance, stress through a reduction of task directed effort, and automation through a reduction in task workload. In accord with the effort regulation hypothesis, then, performance should be poorest under conditions that facilitate a misperception of the effort required for performance and/or which minimize task directed effort.

It is important to note that, although these two performance theories are not mutually exclusive, they often predict substantially different outcomes. Support for one theory, however, does not invalidate the other. Such findings may instead indicate the conditions under which one theory is dominant, and vice versa.

### *Individual Differences in Stress and Performance*

Investigation of individual differences in driver stress requires a distinction to be made between *traits* and *states*. According to Matthews et al. (2000), a trait is "a stable disposition affecting a variety of psychological functions, and a component of personality," while a state is "a more transient reaction [to an event or situation]." In the driving context, traits refer to personality characteristics that moderate vulnerability to stress during vehicle driving, whereas states describe the immediate experience of stress during driving.

*Personality traits.* The Driver Stress Inventory (DSI: Matthews, Desmond, Joyner, & Carcary, 1997) assesses traits related to stress vulnerability. The DSI is a measure of typical emotional reactions to driving in demanding driving conditions, which has been validated in experimental and on-road studies. It assesses five personality traits associated with driving, labeled aggression, dislike of driving, fatigue proneness, hazard monitoring, and thrill seeking. Aggression relates to anger while driving, and to goals for maintaining mastery over other drivers. Dislike of driving corresponds to anxiety and stress associated with the driving task, and to fears about driving competence. Fatigue proneness relates to goals for avoiding discomfort while driving. Hazard monitoring reflects a desire to preempt threat through vigilant search for danger. Finally, thrill seeking relates to an interest in or enjoyment of danger while driving.



The five DSI traits have been shown to be predictive of indicators of real-life driving performance. Aggression, thrill seeking, and less reliably, low hazard monitoring have all been linked to self-reported accident involvement. Matthews (2002) reported that aggression, thrill seeking, and low dislike of driving relate to convictions for driving offenses, such as speeding, and to higher self-reported violations of the law while driving. In addition, aggression, thrill seeking, dislike of driving, fatigue proneness, and low hazard monitoring have all been linked to high rates of self-reported unintentional errors committed while driving (Matthews, 2002).

*Transient states.* Performance and information-processing may also be sensitive to short-lived states such as negative mood, tiredness and worry. Multidimensional assessment of stress states is provided by the Dundee Stress State Questionnaire (DSSQ; Matthews, Joyner, et al., 1999; Matthews et al., 2002). The DSSQ was designed to assess multiple transient states associated with stress, arousal, and fatigue, and to reflect the multidimensionality of these states. It includes 11 scales, which can be grouped into three broad stress dimensions of worry, task engagement, and distress. Several studies reveal that different stressors elicit qualitatively different patterns of stress response, which may have differing impacts on performance. For example, high workload vigilance tasks are associated with reduced task engagement and elevated distress. By contrast, working memory tasks that impose high time pressure are associated with increased distress, increased task engagement, and decreased worry (Matthews, Campbell, et al., 1999). Studies of simulated driving have also shown qualitatively different patterns of subjective response to stressors. For example, driver fatigue elicits task disengagement and distress (similar to vigilance), whereas loss-of-control produces elevated distress and worry (Matthews, 2002).

Studies using the DSSQ confirm that state measures taken prior to performance predict performance on various tasks including vigilance and working memory (Matthews, Campbell, et al., 1999; Matthews, Warm, Dember, Mizoguchi, & Smith, 2001). Support for a resource theory interpretation of such data has come from studies measuring subjective energy, a component of the task engagement factor measured by the DSSQ. Subjective energy reliably predicts performance on sustained attention and 'controlled' search tasks (see Matthews & Davies, 1998, for a review). This relationship has also been demonstrated in studies that show facilitative effects of subjective energy on difficult, attentionally demanding tasks through the construction of Performance Operating Characteristics (POC) for high and low energy participants (Matthews & Margetts, 1991). Recent research has confirmed that the broader task engagement complex, encompassing motivation and concentration as well as energy, also correlates positively with vigilance (Helton, Warm, Matthews, Corcoran, 2002; Matthews et al., 2001). In the driving context, the loss of task engagement characteristic of fatigue may signal a potentially dangerous diminution of the driver's reservoir of attentional resources.

Effects of driver stress personality traits on performance may in part be *mediated* by state responses to the pressures of driving. For example, the effects of aggression on risk-taking behavior may be mediated by the state of anger (Matthews, 2002). Another relevant trait is dislike of driving. Individuals who are high in dislike of driving report higher levels of tension, anxiety, unhappiness, and worry during both real and simulated drives (Matthews, 2001). They also report more negative opinions of driving in general, and appear to experience more negative reactions to environmental stressors than individuals who are low in this trait (Matthews, 1996). Simulator



studies have confirmed that drivers high in dislike of driving tend to perform more poorly, especially when exposed to stress (Matthews, 1996; Matthews et al., 1998). These effects may reflect elevated levels of distress and worry in these stress-vulnerable drivers, although the mediation hypothesis has not been tested directly.

### *Aims of Study*

The purpose for the study was (1) to examine the effects of stress and vehicle automation on driver performance, and (2) to test which trait and state factors were diagnostic of vulnerability to performance impairment. Drivers were allocated to either a nonstressful control condition, or exposed to a stress-induction manipulation intended to elicit a state of distress. Automation was independently manipulated, differing in the effort necessary to regulate speed, resulting in three vehicle automation conditions. In the *automated-driving* condition, speed was set to a fixed value, controlled by the simulator. In the *vehicle-following* condition, the driver was required to follow a lead vehicle driving at constant speed. In the *free-driving* condition, the driver was requested to drive at a constant speed, but there were no other vehicles in the driver's lane, so that speed had to be actively regulated.

Task performance was assessed following the stress induction (or control drive). In addition to driving, participants were also required to detect target stimuli, moving pedestrians, placed among static pedestrian figures (nontargets). Use of a signal detection task allows a comparison of results from this study with those requiring only signal detection.

Predictions were made as follows. Resource theory argues that performance will be poorest when demands on information processing are greatest; stress and workload will have a synergistic effect. Consequently, resource theory predicts performance should be poorest under dual-task conditions in the stress induction free-driving condition. To the extent that vehicle automation successfully reduces task workload, post-task ratings of distress and workload should be correspondingly reduced.

On the other hand, the effort regulation hypothesis predicts that circumstances that reduce active regulation of performance should result in poor task performance. High workload will tend to antagonize stressor effects. Therefore, if the effort regulation hypothesis is correct, performance should be poorest under single-task conditions in the stress induction automated-driving condition. Both dual-task performance and the free-driving condition should elicit higher levels of effort that may protect the driver from stress-related impairment.

Matthews et al. (1997) report that the five factors assessed by the DSI reflect typical subjective emotional reactions to driving and behavior in demanding driving conditions. Previous experimental studies of the DSI factors indicate that they are linked to post-task subjective state and to objective indices of task performance (Matthews et al., 1998). In the present context, it was expected that dislike of driving would relate both to disturbance of subjective state and impairment of task performance. Additionally, it was expected that states of task engagement, as assessed by the DSSQ, would be associated with superior driver performance.



## Method

### *Participants*

A total of 168 students from the University of Cincinnati (69 men, 99 women) participated in the study in order to fulfill a course requirement. Participants ranged in age from 18 to 50 ( $M = 20.37$  years,  $SD = 4.15$ ). All participants had a valid driver's license for a mean duration of 4.95 years.

### *Experimental Design*

A 2 x 3 between groups design was employed in this study, with 28 participants in each condition. Between subjects factors were stress (induction and no-induction), and level of vehicle automation (free-driving, lead-following, and automated-driving). Dependent variables included subjective measures of post-task state and workload, assessed by the Dundee Stress State Questionnaire (DSSQ) (Matthews, Joyner, et al., 1999), and objective measures of task performance obtained through the driving simulator.

### *Questionnaires*

*Driver Stress Inventory. (DSI)* The DSI (Matthews et al., 1997) is an experimentally validated questionnaire designed to assess an individual's vulnerability to stress in a driving context, and to evaluate the coping methods typically employed in stressful driving situations. It is comprised of 41 items designed to assess a driver on five dimensions of driver stress vulnerability: aggression, dislike of driving, hazard monitoring, thrill seeking, and fatigue proneness. DSI scores are scaled so that they may range from 0-100.

*Dundee Stress State Questionnaire (DSSQ).* The DSSQ (Matthews, Joyner, et al., 1999; Matthews et al., 2002) is a 96-item, experimentally validated measure designed to assess transient states associated with stress, arousal, and fatigue, and to reflect the multidimensionality of these states. The items have been factor analyzed into 11 first-order dimensions representing energetic arousal, tense arousal, hedonic tone (pleasantness of mood), intrinsic task motivation, success-related motivation, self-focused attention, self esteem, concentration, confidence and control, task relevant cognitive interference (worry about task performance), and task irrelevant cognitive interference (worry about personal concerns). These correlated scales cluster into three second-order 'super factors' associated with task engagement (energetic arousal, motivation, and concentration), distress (tense arousal, hedonic tone, and confidence and control) and worry (self-focused attention, self-esteem, and both cognitive interference scales). Factor scores for the three higher-order factors were estimated from regression equations that used weights derived from a previous study providing normative data (Matthews et al., 2002). Factor scores are distributed with a mean of 0 and standard deviation of 1, so that the values calculated for a sample represent deviations from normative values in standard deviation units.

The workload items of the post-task DSSQ are six 10-point rating scales reflecting the degree of mental demand, temporal demand, physical demand, performance, effort, and frustration associated



with a task. The scales are modified from the NASA-Task Load Index (NASA-TLX, Hart & Staveland, 1988), a standard measure of workload that is widely used in human performance research (Wickens & Hollands, 2000). The sum of the ratings given to all of the scales provides an overall or global index of workload.

### *Driving Simulator and Performance Tasks*

The simulator used was a Systems Technology, Inc., STISIM Model 400 simulator. The simulator was fully programmable, allowing it to be used to measure performance in a variety of driving contexts. The traffic scene for the simulated drive was projected onto a 38 inch NEC XM3760 monitor. The simulator was equipped with full sized gas and brake pedals and a steering wheel capable of 360 degree steering, which supplied realistic resistance by means of a computer controlled torque motor. Participants completed three driving scenarios in succession, referred to as practice, induction and test phases.

*Practice phase.* In all conditions, participants first completed a fifteen minute practice drive, designed to allow them to become accustomed to the simulator and to practice some of the driving tasks they would encounter later. The practice drive was divided into three periods. First, participants were asked to drive normally, at a speed of 35 M.P.H., to familiarize themselves with the driving simulator. Second, participants practiced driving in the automated condition to which they were assigned: driving normally (free-driving), following a lead vehicle, or steering the vehicle with speed automatically controlled by the computer. Third, participants practiced a hazard monitoring task. During this period, pedestrian pairs were lined up on both sides of the roadway. They were stationed seven feet from the roadway, and spaced 103 feet apart. Occasionally, one pedestrian of a pair would begin to walk into the roadway. The participant's task was to detect these potential road hazards (no maneuvering was necessary to avoid the pedestrian) and to identify which side of the roadway they were on by indicating right on the turn signal for pedestrians on the right side of the road, or left on the turn signal for pedestrians on the left side of the road. The moving pedestrians (the critical signals) were either easy or hard to detect, based on the speed that they were moving (0.5 vs. 0.278 feet per second). At a driving speed of 35 M.P.H., pedestrian pairs (one on the left and one on the right sides of the roadway) constituted a stimulus event rate of 30/minute. Participants had six seconds to detect and respond to the moving pedestrians (five seconds before the participant passed the pedestrian, and one second after they passed the participant). The computer recorded the participant's reaction time and response to each moving pedestrian (correct, incorrect, and no response).

*Induction phase.* Following practice, participants drove manually through an "induction" phase regardless of the automation condition to which they were assigned. In both the stress-induction and no-stress-induction conditions, the induction phase took 20 minutes to complete. In the stress-induction condition, participants were warned that they might encounter hazardous conditions similar to "black ice". Participants were allowed to drive normally for the first 2.5 minutes of the drive. After this period of time, they were occasionally subjected to unpredictable periods of loss-of-vehicle-control. This was accomplished by reducing the coefficient of friction of the roadway, and by introducing a heavy wind gust into the driving scenario. The coefficient of friction of the road surface was reduced from its default value of 8 to a value of 1. The wind gusts



employed were constructed from a single sine wave, with amplitude of 1000 and frequency of 1 peak every 12 seconds. For the simulator, the amplitude of a wind gust is specified as a true lateral wind speed in feet/second, meaning that the participant experienced the wind gusts as blowing laterally at speeds of up to 1000 feet/second. The wind gusts were randomly assigned to 10, 15, 20, or 25 seconds in length. Wind gusts were positioned randomly throughout the drive, with one taking place approximately every 48 seconds. The net effect of these manipulations was that the participant's vehicle was pushed uncontrollably around the roadway. In the no stress induction condition, participants were also warned that they might encounter conditions of black ice, but they were never exposed to any loss of vehicle control.

*Test phase.* At the beginning of the test phase, participants resumed driving in the automated condition to which they were assigned. Drivers in the automated-driving condition had their vehicle speed set to 35 M.P.H. Drivers in the lead-following condition were required to follow a lead vehicle. Drivers in the free-driving condition were instructed to drive at a constant 35 M.P.H. The initial part of the drive was a period of approximately four minutes of normal, single-task driving, followed by the dual task of driving combined with monitoring for pedestrian hazards. During the hazard monitoring task, the pedestrian event rate was 30 pairs of pedestrians per minute. There were 56 critical signals during the test phase, occurring at a rate of 4 critical signals per minute. The hazard monitoring task continued for sixteen minutes (at 35 M.P.H.), and then the simulator terminated the program. The entire test phase took approximately 20 minutes to complete. During the test phase, participants in the stress-induction condition experienced three brief loss-of-control episodes to maintain their state of stress. Indices of performance, including lane position, speed, and turn signal responses to the pedestrian stimuli were logged by the computer throughout the drive.

### *Procedure*

Participants were tested in a windowless laboratory room. When they first arrived, they were required to read and sign an informed consent form, and to present the experimenter with a valid driver's license. Next, participants were assigned to an experimental condition, according to a randomized schedule. They were then asked to complete the Driver Stress Inventory (DSI) and the pre-task Dundee Stress State Questionnaire (DSSQ). After completing these questionnaires, participants were informed that they would begin the driving task of the experiment. Each of the participants was asked to sit in a computer chair approximately 4 feet away from the monitor, at a small table that contained the steering box, and below which the gas and brake pedals were positioned. Participants were directed to position the chair at a height and distance from the wheel and pedals that was comfortable to them. Following instructions, each of the participants performed the practice phase of the drive, followed by the induction and test phases, according to the experimental conditions to which the person was assigned. After the termination of all three phases of the driving task, participants completed the post-task Dundee Stress State Questionnaire (DSSQ). Participants were instructed to complete the DSSQ, including its embedded workload items, with specific reference to the *dual-task* phase of the drive.



## Results

### Subjective Stress

The data for each DSSQ factor were tested for statistical significance by means of a 2 (phase)  $\times$  2 (stress induction)  $\times$  3 (vehicle automation) analysis of variance (ANOVA) in which stress induction and vehicle automation were between-subjects factors and phase was a within-subject factor. The aim of the analysis was to test for changes in state from pre- to post-task (indicated by a main effect of phase) and for variation of such changes with the experimental factors (indicated by phase  $\times$  factor interactions). The analyses of the data for the worry and task engagement factors revealed that the phase difference was statistically significant for each factor,  $F_{\text{worry}}(1, 62) = 56.47, p < .01$ ;  $F_{\text{task engagement}}(1, 62) = 157.46, p < .01$ . All of the remaining sources of variance in the analyses of the worry and task engagement factors lacked significance ( $p > .05$ ). The phase difference was also statistically significant in the case of the distress factor,  $F(1, 62) = 281.74, p < .01$ . For this factor, however the phase difference was also dependent upon both stress induction,  $F_{\text{phase} \times \text{stress induction}}(1, 62) = 11.77, p < .01$ , and automation,  $F_{\text{phase} \times \text{automation}}(2, 162) = 4.22, p < .05$ . All other sources of variance in the analysis of the distress factor were not significant ( $p > .05$ ).

Change scores for the three main DSSQ factors as a function of stress condition and vehicle automation are shown in Figure 8. The Figure shows that, in all conditions, participants' levels of worry and task engagement were lower at the end of the experimental session than prior to its start, while their level of distress increased from pre-test to post-test. The effects of phase on distress were moderated both by the stress induction and by automation. Figure 8 shows that the increment in distress was greater for participants who experienced the stress-induction procedure than for those who did not experience that procedure. In addition, the distress response was smaller in the automation and following driving conditions than it was in the free driving condition.

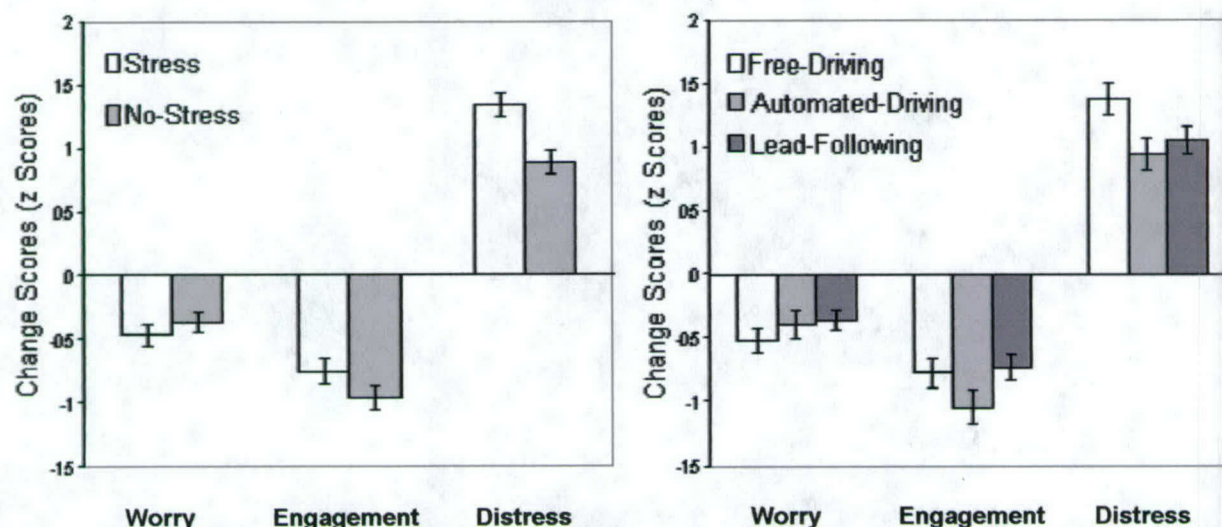


Figure 8. Change scores for the three DSSQ factors as a function of stress condition (left panel) and vehicle automation condition (right panel).



### *Perceived Mental Workload*

Mean overall workload scores for the dual-task phase of the drive and their associated standard errors are presented for all experimental conditions in Table 4.

Table 4. Mean global workload scores under all experimental conditions. Values in parentheses are standard errors.

Stress Induction	Automation	Workload
Induction	Free-driving	35.93 (1.37)
	Automated-driving	34.46 (1.95)
	Lead-following	37.38 (1.20)
No induction	Free-driving	31.00 (1.72)
	Automated-driving	29.46 (1.60)
	Lead-following	30.43 (1.79)

A 2 (stress induction)  $\times$  3 (vehicle automation) ANOVA revealed a statistically significant main effect for stress induction,  $F(1, 162) = 18.03, p < .01$ . Participants who experienced stress induction reported higher workload for the dual task phase of the drive ( $M = 35.92, SE = .89$ ) than those who did not experience stress induction ( $M = 30.30, SE = .97$ ). The effects of vehicle automation and the interaction of stress induction and vehicle automation were not statistically significant ( $p > .05$ ).

### *Performance Variables*

During the test phase of the drive, the simulator recorded data at a rate of ten samples per second. The principal performance indices were standard deviation in lane position (a commonly-used index of psychomotor control) and the frequency of correct detections on the signal detection task. Also monitored was vehicle speed in the non-automated conditions, primarily to check for tradeoffs between the other performance indices and speed. In fact, no such tradeoffs were found, and therefore, speed data are not reported here. To track changes in performance across the test phase, the data were separated into four segments. Segment one, corresponding to minute 4 of the test phase, was selected as a baseline measure of single-task performance. Segments two through four, corresponding to the mean of minutes 6 and 7, minutes 12 and 13, and minutes 17 and 18, respectively, were selected as representatives of initial, intermediate, and concluding dual-task performance.

*SD of Lane Position.* A 4 (segment)  $\times$  2 (stress induction)  $\times$  3 (vehicle automation) mixed model ANOVA was calculated for the standard deviation in lane position in which stress induction and vehicle automation were between-subjects factors and segment a within-subject factor. In this and all subsequent analyses involving repeated measures with more than two levels of the factor, Box's epsilon was employed to correct for violations of the sphericity assumption (Maxwell & Delaney, 2004).



For the standard deviation in lane position, statistically significant main effects were found for segment,  $F(2.67, 430.31) = 20.56, p < .01$ , and automation,  $F(2, 162) = 3.76, p < .05$ . All other sources of variance were not significant ( $p > .05$ ). Figure 9 shows that SD of lane position was higher during dual-task performance (segments 2-4) than in single-task performance (segment 1). The Figure also shows that participants in the free-driving condition displayed greater lane deviations than participants in the automated-driving and lead-following conditions.

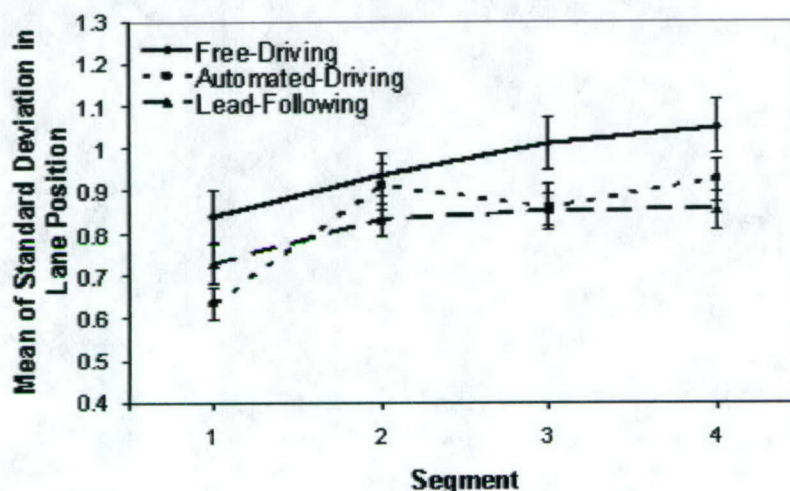


Figure 9. Mean of standard deviation in lane position for each vehicle-automation condition as a function of segment. Error bars are standard errors.

*Secondary Task Detections.* The simulator also recorded the number of correct easy and hard secondary task detections. A 2 (stress induction)  $\times$  3 (vehicle automation) between-subject ANOVAs was calculated separately for each type of secondary target. The mean number of correct secondary task detections and associated standard errors for each of the vehicle automation conditions can be seen in Figure 10. For both easy and hard correct detections, statistically significant main effects of vehicle automation were found,  $F_s(2, 162) = 11.53, p < .01$ , and 3.29,  $p < .05$ , respectively. Participants in the free-driving condition detected the fewest signals in both cases, as shown in Figure 10.



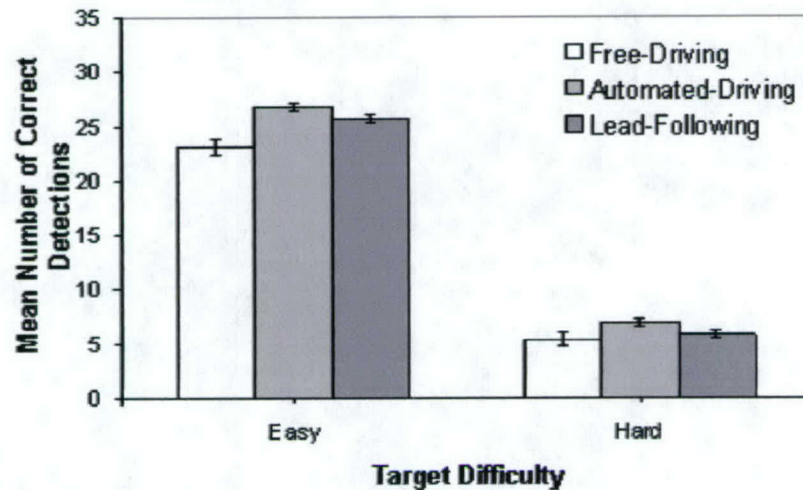


Figure 10. Mean number of correct secondary task detections for each vehicle automation condition and target difficulty. Error bars are standard errors.

### Individual Differences

*Predictors of Post-Task State.* Table 5 displays the correlations between the pre-task and post-task DSSQ secondary factors and the five dimensions of driver stress vulnerability assessed by the DSI. These data confirm that subjective state is meaningfully related to stable personality factors.

Table 5. Correlations between DSSQ pre- and post-task scores and DSI factors.

DSI Factors	Worry		Task engagement		Distress	
	Pre	Post	Pre	Post	Pre	Post
Aggression	.31**	.26**	-.24**	-.24**	.20**	.09
Dislike of driving	.28**	.29**	-.23**	-.08	.49**	.38**
Hazard monitoring	.11	.09	.31**	.19*	-.21**	-.22**
Thrill seeking	.21**	.17*	.02	-.06	-.11	-.13
Fatigue proneness	.19*	.23**	-.29**	-.16*	.40**	.26**

Note: \*  $p < .05$ , \*\*  $p < .01$

*Predictors of Task Performance.* To assess the impact of DSI (trait) and DSSQ (state) factors on dual-task performance, hierarchical multiple regressions with four steps were conducted. First, the means of the two key performance variables recorded by the simulator (the SD in lane position, and total number of detections) during dual-task driving were calculated. These performance indices were then individually used as criteria in two hierarchical multiple regressions. The four steps, entered successively, were effect-coded vectors for stress-induction



and vehicle-automation conditions, effect-coded vectors for the interaction of stress induction and vehicle automation, post-task DSSQ scores, and the DSI factors.

Table 6. Regression of the standard deviation in lane position onto task, state and trait variables.

Step	Variables entered	df	$R^2$	$\Delta R^2$	F for $\Delta$	Statistically significant predictors
1	Stress condition Automated condition	3, 164	.042	.042	2.42	
2	Product vector of stress and automation	2, 162	.045	.002	.21	
3	Post-task DSSQ scores	3, 159	.136	.091	5.57	Post-task worry ( $\beta = .210^{**}$ ) Post-task task engagement ( $\beta = -.220^{**}$ )
4	DSI factors	5, 154	.246	.110	4.51	Dislike of driving ( $\beta = .283^{**}$ ) Thrill seeking ( $\beta = -.255^{**}$ ) Fatigue proneness ( $\beta = .278^{**}$ )

Note: \*  $p < .05$ , \*\*  $p < .01$

Summary statistics are given in Table 6. The task variables were predictive as in the ANOVAs, so here the focus will be on the extent to which the individual difference factors added to the variance explained by task factors. Statistically significant predictors were worry and task engagement, and the DSI factors of dislike of driving, thrill seeking, and fatigue proneness ( $F [13, 154] = 3.87, p < .01$ ). The inclusion of the DSSQ post-task factors (step 3) and the DSI factors (step 4) added over 20% to the variance explained by the model. For the final model, standardized betas for post-task worry, post-task task engagement, dislike of driving, thrill seeking, and fatigue proneness were .21, -.22, -.28, -.26, and .28 respectively (all  $p < .01$ ).

No significant DSI or DSSQ factors were found to be predictive of number of correct secondary task detections. Also conducted were tests of whether the diagnosticity of DSSQ factors in the prediction of vehicle control varied with stress induction condition, by testing whether product terms representing interactions between the stress induction and the three DSSQ factors added significantly to the variance explained in the regressions. In fact, the interaction terms were non-significant, implying that high worry and low engagement were equally diagnostic of poorer vehicle control in both stress and non-stress conditions.

### Discussion

In this study, the effects of stress induction and vehicle automation on driver stress and performance were explored. Consistent with the transactional model of stress, the results confirmed that driver stress and performance are sensitive to stress induction and vehicle automation as well as to personality traits related to stress vulnerability.

*Effects of experimental manipulations.* In general, the results provided only partial support for the initial hypotheses. As predicted, the stress induction had a statistically significant



effect on both post-task distress and on subjective workload, indicating that the stress induction was successful. However, the stress induction had little effect on objective task performance variables. The influence of vehicle automation, compared with free driving, showed effects consistent with expectation in reducing post-drive distress and enhancing driver performance, including vehicle control (lower SD of lane position) and attention (detection of secondary task stimuli). Contrary to expectation, automation failed to reduce workload, and failed to interact with the stress induction in its effects on subjective and objective measures.

The theoretical context for the present study was that both resource theory (Wickens & Hollands, 2000) and effort-regulation theory (Hancock & Warm, 1989) appear promising for explaining stress and automation effects on driver behavior. In fact, findings were more supportive of predictions from resource theory than from the effort-regulation hypothesis. Consistent with previous studies of dual-task interference while driving (Kantowitz & Simsek, 2001), the transition from single- to dual-task performance had a substantial negative impact on driver performance. Furthermore, subjective workload was high, implying substantial allocation of resources (Wickens & Hollands, 2000), and, therefore, vulnerability to resource shortfalls.

Assessment of the objective performance indices obtained in this experiment indicated that vehicle automation had a facilitative effect on driver performance. As the driver became responsible for fewer components of the driving task, performance on the standard deviation in lane position, longitudinal speed, and the number of correct secondary task detections was improved. Consistently, the free-driving condition exhibited the poorest driver performance. However, fewer performance differences between vehicle-following and automated-driving were found, suggesting that the benefits of automation are somewhat limited. These findings support resource theory because driver performance was influenced by the afforded degree of autonomy over speed. As the driver's direct responsibility for speed regulation diminished, driver performance was concomitantly improved.

Perhaps surprisingly, the stress induction had little effect on performance. A resource-based explanation may be found in other simulator studies (e.g., Matthews, 1996) which suggest that the effects of stress induction are most pronounced when task demands are low (i.e., on straight sections of roadway, rather than on curved sections). In this experiment, however, high dual-task difficulty may have forced drivers to maintain task-directed effort, as also suggested by the reduced level of post-task worry.

*Individual differences.* Relationships between personality, indexed by the DSI, and subjective state were generally consistent with previous findings (Matthews, 2002). Consistent with the transactional model of stress, subjective distress reflected personality as well as external stress factors. Specifically, post-task distress was correlated with high DSI dislike of driving, and was sensitive to the stress induction and automation manipulations.

The regression analyses showed that vehicle control (SD of lane position) could be predicted from the trait and state variables, but secondary task detections could not. In fact, there were multiple independent predictors of SD of lane position, including both trait factors (e.g., high dislike of driving, fatigue proneness) and state factors (high worry, low engagement), that, taken together, added about 20% to the variance in performance explained. The facilitative effects



of task engagement (as a measure overlapping with subjective energy) detected in this study underscore the dependent relationship between task engagement and difficult, attentionally demanding tasks, which has been demonstrated in other studies (Matthews & Davies, 1998). This finding provides further evidence that performance may have been controlled by resource availability rather than effort-regulation in this particular study.

It is perhaps a little surprising that task engagement (and other predictors) was linked to psychomotor control rather than signal detection, given the robust finding from other studies that task engagement and energy relate to vigilance (Matthews & Davies, 1998; Matthews et al., 2001). One possible explanation comes from a prior dual-task study in which subjects performed two visual search tasks concurrently (Matthews & Margetts, 1991). Energetic arousal was related to speed of search only of the higher priority search, and was unrelated to speed of the lower priority search. It was suggested that high energy confers additional resources that, in line with effects of arousal on attentional selectivity, are allocated to the primary task activity. The greater sensitivity of vehicle steering to individual difference factors in this study may reflect its status as the more important of the two tasks performed.

*Conclusions and Human Factors Implications.* Broadly, the results of this experiment support a resource-based model of driver performance. However, the difficulty of the dual-component driving task employed in this experiment may have mitigated effects of task variables and personality factors believed to control performance in low workload conditions, and, therefore, the findings of this experiment must be interpreted with some caution. The observed facilitative effects of vehicle automation in this experiment offer some hope that future in-vehicle automation will indeed aid driver performance, either through reducing driver workload, or, as was suggested here, by allowing drivers to reallocate resources to other task demands. It must be noted, though, that the automation employed in this experiment was involuntary. That is, participants had no control over either the type of automation they experienced or when the automation was activated. As Parasuraman and Riley (1997) point out, deliberations over when to activate automation may impose additional cognitive workload over and above that already required by a task.

Both stress and automation are potentially human factors problems, but the data from this experiment indicates that they may operate independently rather than synergistically. Thus, the present data may somewhat reassure designers of automated systems that they do not need to focus excessively on stress-related issues. However, this conclusion is highly tentative because of the lack of other studies examining the effects of driver stress and vehicle automation. Further, there are concerns that poorly-designed automation may itself be a source of stress (Parasuraman & Riley, 1997), and it is still possible that automation and stress may interact in underload conditions, rather than the high workload conditions imposed in the present study (as evidenced by workload data). For example, automation may add to operator vulnerability to stress during monotonous convoy operations.

This study also demonstrates that individual differences in driver performance may be predicted from trait and state measures. As in previous studies (e.g., Matthews et al., 1998), variability in lateral position appears to be an especially sensitive measure of performance. Since a high proportion of army vehicle accidents involve vehicle roll-over, the driver's ability to steer



accurately is also relevant. The data show that maximizing prediction of performance requires use of *multiple* predictors. No single state or trait predictor in this study explained more than 10% of the variance in performance, but, together, they added over 20% to the criterion variance explained by task factors. A detailed profile of the individual's vulnerability to stress and actual stress experienced is necessary for optimal prediction.

Trait factors are most relevant to the domain of driver personnel selection, where employees must be selected based on their driving competence and on their safety records. The personality traits measured by the DSI may be particularly pertinent to this selection process, as those traits represent stable, dispositional characteristics that may substantially impact driver performance, particularly in stressful driving situations (Matthews, 2001). The usefulness of such measures has already been demonstrated by studies showing that the DSI is predictive of real-world frequencies of accident involvement and of convictions for driving offences (e.g. Matthews et al., 1997). By incorporating measures such as the DSI into employee screening processes, safe candidates may be more easily selected, while less suitable employees are screened out. Personality measures such as those provided by the DSI may also be used to identify drivers who have special needs for training or remedial programs because of traits such as high dislike of driving or aggression.

In predicting safety during a specific drive requiring sustained performance, the data show that trait measures must be supplemented with indices of state, as assessed by the DSSQ. Specifically, monitoring for high levels of task disengagement and worry may be important for diagnosing the at-risk driver. In the military context, the risks of disengagement correspond to the well-known risks associated with fatigue. However, worry has perhaps been insufficiently studied as a risk factor. As noted previously, the demands of wartime convoy driving, such as the ever-present risk of sniper or bomb attacks, may be conducive to the development of worry in some individuals.

In sum, the implications of the current study for human factors researchers are clear. As this experiment is one of very few that have directly examined the effects of subjective stress *and* vehicle automation on driver performance, further research incorporating both factors is necessary. Stressful driving conditions may be experienced by drivers on a daily basis. Yet, the potential interaction of stress and automation has not been thoroughly investigated, even as in-vehicle technology increases at a rapid pace. The present research suggests that future research may be able to identify diagnostic predictors of risk during sustained driving, predictors that generalize across nonstressful and stressful driving conditions.



### STUDY 3: DIAGNOSTIC METHODS FOR PREDICTING PERFORMANCE IMPAIRMENT DURING A SUSTAINED ATTENTION TASK

#### Disclaimer

**This study is still in progress. Thus far, data have been collected from 97 participants, but data collection will continue into the second year of the project. Initial findings are presented here, but it is emphasized that these are of a preliminary nature. Note also that salivary cortisol data are not yet available. A complete report on this study will be provided in the second annual report on this project.**

**The authors of this report request that this part of the document is not circulated to any persons other than those directly involved in assessment of the report.**

Stress experienced by military personnel may take many forms, including physical challenges such as noise and heat, task demands such as cognitive overload and time pressure, and psychological stress factors such as anxiety. Traditionally, problems of this kind have been addressed within a 'stressor-strain' model. Stress factors (*'stressors'*) produce strain within the individual, and if strain is excessive, performance is impaired. Increasingly, evidence suggests that the simple stressor-strain model is inadequate. It is often difficult to predict the impact of stress factors on performance, and there are typically pronounced individual differences in the impact of such stressors on performance.

Recently, it has been argued that a multidimensional approach is required (Matthews, 2001). It appears that stressors may produce multiple, independent changes in physiological and psychological functioning, changes that cannot be adequately described in terms of general stress or arousal. Relevant physiological changes – often poorly intercorrelated with one another (Thayer, 1989) – include autonomic (a.n.s.) and central nervous system (c.n.s.) responses and hormone secretion. At the psychological level, our recent work (Matthews et al., 2002) discriminates three qualitatively different stress states - task disengagement (related to fatigue), distress and worry - through psychometric, experimental and correlational studies. There is some overlap between subjective states and psychophysiology; for example, the DSSQ correlates with various a.n.s. and c.n.s. indicators (Venables & Fairclough, 2004). However, neither type of index is fully predictable from the other (Thayer, 1989), so that a complete assessment of the state of the individual requires both subjective and physiological measures.

The relation between different dimensions of the stress state (e.g., subjective anxiety, a.n.s. arousal) and performance is also multi-faceted. The tasks performed by military personnel impose differing cognitive demands, ranging from basic psychomotor skills to high-level tactical decision-making. Abundant evidence suggests that the relation between stress factors and performance depends critically on the specific information-processing components that support task performance (Hockey, 1986; Matthews et al., 2000). Our current work is beginning to establish mappings between specific stress states and specific types of task. Research so far links task disengagement to impaired sustained attention, distress to multiple-task performance, and worry to high-level verbal skills (Matthews, 2001). It is likely that physiological indicators are also



specific in regard to the aspects of performance to which they relate; we cannot expect that any single indicator will predict the full spectrum of cognitive tasks.

Two practical conclusions follow from this analysis of stress and performance. First, because of the variability of responses to stress, it is important to assess fitness to perform at the level of the individual soldier. Second, no single index will provide a universally valid diagnosis of fitness to perform. Predicting performance is likely to require information from multiple indicators, both physiological and psychological. Next, we briefly review some of the indicators that may prove useful for predicting sustained attention, given that traditional a.n.s. measures such as heart rate and skin conductance do not appear to be good predictors (Davies & Parasuraman, 1982).

### *Transcranial Doppler Sonography*

As previously discussed (see Study 1, above), studies at UC (Mayleben, 1998; Hitchcock et al., 2003; Hollander et al., 2003) show that decreases in blood flow in the cerebral arteries, measured using transcranial Doppler sonography (TCD), are linked to loss of sustained attention. TCD may provide a non-invasive index of cerebral functional status that is promising for efforts to predict performance in stressful environments. However, as with other physiological indices, work using TCD has primarily explored how task parameters may control both physiological and behavioral response. Studies at UC suggest that TCD may be useful for monitoring the development of functional impairment *during* sustained performance: decline in right hemisphere blood flow parallels loss of perceptual sensitivity. However, it is unclear whether TCD can be used to predict loss of performance *in advance of* performance. Thus, the study will test whether a short-duration blood flow assessment can be used to predict later performance impairment, i.e., to develop and test a *diagnostic* technique for performance impairment.

### *Salivary cortisol*

Much recent work on the psychophysiology of stress response (reviewed by Dickerson & Kemeny, 2004) has focused on the role of cortisol, a steroid hormone that may index activity of the hypothalamic-pituitary-adrenocortical (HPA) axis (sometimes known as the fight-or-flight response). The HPA has an acute 'proactive' role in mobilizing energy and physiological defense mechanisms, and a longer-term 'protective role' associated with immunosuppression and other processes that prevent damage associated with over-activity of defense mechanisms (see Buckingham, Gillies, & Cowell, 1997 for a review). Studies of combat veterans suffering from PTSD (Mason et al., 2001) suggest that cortisol levels reflect the ongoing balance between the undifferentiated emotional arousal state of engagement (associated with higher cortisol levels) and opposing defense mechanisms leading to deactivation (associated with lower cortisol levels). Magnitude of the cortisol response can be readily assayed from saliva (Kirschbaum & Hellhammer, 2000). Although a variety of stressors provoke cortisol release, relationships between cortisol and subjective measures, such as anxiety, are inconsistent, implying that assessment of cortisol may provide diagnostic information additional to that provided by self-report (Huwe, Henning & Netter, 1998).



The functional significance of cortisol for attention and performance is not well understood, given that in stress research, cortisol is more often studied as a dependent variable than as a criterion variable. Recent work is providing increasing evidence that cortisol levels may correlate with performance, although the relation between cortisol and information-processing is still not well-understood. In general, results suggest that cortisol levels may be associated with both deficits and enhancements in performance, depending on the task. Positive findings included speeding of reaction time (Skoschnik et al., 2000) and enhancement of memory in male subjects (Wolf et al., 2001). During vigilance, high cortisol correlates with electroencephalographic indices of stress and alertness (Born et al., 1988). On the other hand, high levels of cortisol may be indicative of impairments in selective attention (Lupien et al., 1994; Wolf et al., 2001). It is likely that use of cortisol as a diagnostic index requires attention to the temporal dynamics of the response. For example, Sluiter et al. (2001), in a workplace study, found that cortisol reactivity during work was correlated with an increased need for recovery after work. Somewhat similarly, elevated cortisol response during sustained cognitive performance is correlated with impaired attention following the period of work (Bohnen et al., 1990). Systematic performance-focused research in this area is urgently needed.

### *Subjective measures*

The principal subjective measure employed is the DSSQ (Matthews, Joyner, et al., 1999; Matthews et al. 2002). Psychometric analyses performed within the initial development sample ( $N=767$ ), distinguished 10 inter-correlated factors relating to mood (e.g. tense arousal), cognition (e.g. interfering thoughts) and motivation. A further factor analysis of these dimensions obtained three factors integrating affect, motivation and cognition, that were labeled as Task Engagement, Distress and Worry (see Table 1). Other studies confirmed various psychometric properties required of state measures: high reliability, day-to-day variation of scores, and distinctiveness from trait measures. Arousal dimensions of the DSSQ correlate with psychophysiological indices of autonomic arousal (Matthews, 1987).

Table 7. Three second-order stress state dimensions.

Task Engagement	Distress	Worry
Energetic arousal	Tense arousal	Self-consciousness
Motivation	Low hedonic tone	Low self-esteem
Concentration	Low confidence	Cognitive interference (task worries)
		Cognitive interference (personal concerns)

*Note.* Some primary scales have additional, minor loadings on other secondary factors

Outcomes of studies using the DSSQ, and other published work, suggest the following summary of three distinct modes of stress response (see Matthews et al., 2000; Matthews et al. 2002; Matthews, Joyner, et al., 1999, for detailed reviews):

- *Task disengagement:* feelings of fatigue and tiredness, loss of motivation and distractibility. Prolonged monotonous tasks such as vigilance and sustained vehicle driving tend to elicit task



disengagement. Low task engagement is associated with impairments of controlled processing, vigilance, and more complex tasks requiring sustained attention.

- *Distress*: typified by feelings of tension, unhappiness and frustration, and lack of confidence and control. Distress is elicited especially by cognitive overload derived from task demands and time pressure. Our recent research (e.g., Matthews et al., 1998; Matthews & Campbell, 1999) suggests that distress is particularly damaging to tasks requiring executive control of multiple task components, such as working memory, and concurrent attention and tracking.
- *Worry*: Intrusive thoughts about the task and personal concerns, self-focused attention, and loss of self-esteem. Failure experiences seem to be a potent source of worry, which especially harms complex, verbally-mediated tasks such as reasoning and verbal communication.

The task engagement dimension appears to be the aspect of state most likely to predict sustained attention. As previously discussed (see Study 2), numerous studies have shown that one of its core components, energetic arousal, predicts performance on vigilance and other demanding attentional tasks (Matthews & Davies, 1998). Recent work has also shown that both task engagement and energetic arousal correlate with performance on vigilance tasks requiring judgements of line length (Matthews, Campbell, et al., 1999) and discrimination of single characters (Helton et al., 2002; Matthews et al., 2001). These studies also showed a weaker tendency for high distress and worry to relate to poorer vigilance.

In addition to the DSSQ, we included two further subjective measures as criterion rather than predictor variables. The first such measure is a modified version of the NASA-TLX (Hart & Staveland, 1988), a widely used scale for assessing workload. Assessment of workload following performance of the sustained attention task affords (1) confirmation that the task does indeed impose a substantial workload, and (2) a test for congruence of predictors of objective performance with subjective workload. The second criterion measure is a measure of situation awareness, referring to the person's perception and comprehension of the task situation. Maintaining situation awareness has been seen as especially important for soldiers, given the complex and dynamic nature of combat (e.g., Endsley, 1995).

### *Aims and hypotheses*

This study was designed to use indices of physiological and subjective state to predict performance during a vigilance task. It is intended that it will provide comparative data on the validity of psychophysiological indices of self-report state, salivary cortisol, and cerebral blood flow (CBF), measured by transcranial Doppler sonography (TCD), as indicators of performance and functional efficiency. Because little is known about individual differences in blood flow, it is also necessary to investigate psychometric properties, such as reliability of blood flow indices.

This study examines both concurrent and predictive utility of the psychophysiological indices. First, it designed to test whether psychophysiological response to a brief cognitive challenge is predictive of *future* performance impairment. Participants complete a short battery of highly demanding tasks, validated as high-workload tasks in our laboratory. The tasks are expected to be sufficiently stressful to elicit subjective stress and salivary cortisol responses. Tasks



similar to those employed in this study have been shown to elicit increases in blood flow over short durations (Stroobant & Vingerhoets, 2000). We hypothesized that the magnitude of the increase in blood flow will be predictive of subsequent vigilance performance.

Second, the study will also test whether TCD may be used as a *concurrent* index of loss of functional efficiency during vigilance. Specifically, we to determine whether blood flow measured during vigilance is correlated with detection rate. Generally, it is hypothesized that optimal prediction requires a multivariate approach of using multiple predictors, including subjective state and cortisol measures in addition to blood flow.

### Method

*Participants.* Thus far, 41 men and 56 women have been recruited from the Introductory Psychology subject pool at the University of Cincinnati. All meet the following exclusion criteria:

- Impaired vision (corrected vision is acceptable)
- Currently taking psychoactive medication
- Currently taking medication containing corticosteroids
- Physically unable to perform the cognitive tasks
- History of epilepsy
- Lack of fluency in English

Participants are also asked to meet the following requirements at the time of testing (to prevent extraneous influences on the cortisol assays):

- Must refrain from alcohol use for 24 hours prior to participation
- Must refrain from substantial meals, drinks containing citrus, caffeinated products, tobacco use, and vigorous physical exercise for 2 hours prior to participation

### Design

All participants follow the same general protocol. The sequence of events is:

- Completion of personality questionnaire
- Completion of baseline stress state questionnaire; saliva sample collection
- Assessment of baseline cerebral blood flow, using TCD
- Performance of short battery of three demanding tasks
- Completion of post-task stress state questionnaire; saliva sample collection
- Performance of vigilance task
- Completion of post-task stress state and situation awareness questionnaires; saliva sample collection



## Questionnaires

*Personality.* The 40-item Saucier (Saucier & Goldberg, 2002) Mini-Markers questionnaire assesses the 'Big Five' dimensions of personality: Extraversion, Neuroticism, Agreeableness, Conscientiousness and Openness. These are believed to represent fundamental, universal personality traits.

*Stress state.* The DSSQ (Matthews Joyner, et al., 1999; Matthews et al., 2002) assesses participants' immediate moods, motivations, cognitions and coping strategies, prior to or following task performance. It may be scored for three broad subjective state factors, task engagement, distress and worry. The post-task version also includes a short workload assessment, based on the NASA-TLX (Hart & Staveland, 1988). A more detailed description is provided in the Method for Study 2.

*Situation awareness.* The measure used here was a subjective rating system developed by Taylor (1989), the Situational Awareness Rating Technique (SART). The respondent rates 10 aspects of situation awareness during performance, that may be grouped to form three higher order dimensions: Demands on Attentional Resources, Supply of Attentional Resources, and Understanding of Situation.

## Performance tasks

*Short battery of high workload task.* The battery is comprised of three tasks. In each case, stimuli are presented on a computer monitor, and the participant responds using the keyboard or mouse. The first task (*line length discrimination*) is a signal detection task in which the participant is presented with pairs of flickering lines; stimulus duration is 300 ms. Stimuli are presented at a rate of 1/second. The critical signal is a pair of lines longer than the standard length. On detection of a critical signal, the participant is required to press the spacebar. The second task requires *working memory*. The participant views a series of math problems and presses a key to indicate whether or not the problem is correct. Above each problem is a concrete noun. After each series of problems, the participant must record the first or last letter of each noun, in serial order. The third task is a *tracking* task. A target moves unpredictably according to a forcing function based on a sum of sine functions. A cursor is also presented on the screen; the participant uses a joystick to track the target with the cursor as accurately as possible. All three tasks are designed to impose a high workload.

*Vigilance task* (Hitchcock et al., 1999; Hitchcock et al., 2003). The vigilance task requires participants to monitor a simulated air-traffic control display. A computer is used to present stimuli on a 17" monitor and record all participants' responses. Critical signals for detection (emergency events) are cases in which two aircraft, one flying on a northwest/southeast heading and the other flying on a northeast/southwest heading are aligned on a collision path over the center of a city. Participants indicate their detection of a critical signal by pressing the space bar on a keyboard. Intervals between critical signals ranged from 12 to 60 sec with a mean of 30 sec. In all conditions, the display is updated 30 times/minute with a stimulus presentation time of 300



ms. Hitchcock et al. (2003) showed that the task imposes a high workload and produces temporal decrements in both performance and cerebral blood flow.

### *Procedure*

Participants are tested in a windowless laboratory room. Upon arrival, they are required to read and sign an informed consent form, following an informed consent interview, and to confirm that they meet exclusion criteria. Participants then complete the Saucier Mini-Markers and pre-task DSSQ. They also provide a saliva sample by chewing on a cotton wool swab for c. 2 minutes to produce saliva. The swab is placed in a plastic container and refrigerated until shipped to an external company (Salimetrics) for a cortisol assay. Next, the TCD ultrasound transducer is secured around the head; see the Method of Study 1 for a complete description of the procedure. By contrast with Study 1, two transducers are fitted within the head bracket in order to record blood flow bilaterally, using a Nicolet Companion III unit. A baseline blood flow measure is then taken for five minutes, while the participant views a blank screen with no performance imperative. Next, the participant performs the three short tasks for a period of 2 minutes each. There is a period of 2 minutes between each task, during which additional baseline measures of blood flow are taken. After performance of the third task, participants complete a post-task DSSQ, and provide a further saliva sample. Because the salivary cortisol response to a stressor is typically delayed by about 20 minutes, an assay taken after the DSSQ should be diagnostic of stress response during the battery of high-workload tasks. The order in which the workload tasks are presented is counterbalanced across participants.

After completing the short task battery, participants practice the vigilance task with feedback provided following errors. They then perform the full 40-minute task, divided into four continuous 10-min periods of watch without feedback. Cerebral blood flow is recorded bilaterally throughout performance of the vigilance task. The principal dependent variable is the participant's accuracy in detecting signals (per cent correct detections). Upon conclusion of the vigilance task, participants complete a further post-task DSSQ, followed by the SART situation awareness measure, and provide another saliva sample. Participants are then debriefed.

## **Results**

The analyses presented here are not exhaustive, and are intended to indicate some principal findings of this ongoing study. A more detailed analysis, including multivariate analyses, will be presented in the next report. As previously indicated, all findings should be viewed as tentative at this point. Results are divided into three sections: manipulation checks, individual differences in blood flow, and initial correlational data.

### *Manipulation Checks*

*Subjective state response.* Previous work (e.g., Matthews et al., 2002) suggests that the short demanding tasks should provoke subjective responses characterized by increased distress. Figure 11 shows change scores compared to baseline for the three DSSQ secondary factors following the short battery (left panel) and following the vigilance task (right panel). In each case, the score represents the difference between post-task subjective state and the initial baseline



measurement, expressed in standard score units. Error bars in these, and all subsequent figures, represent  $\pm 1$  standard error of the mean. Both the short battery and the vigilance task elevated distress and depressed worry. However, the short battery left task engagement unchanged, while the vigilance task induced a large magnitude decline in engagement. These data confirm that both tasks exposures are broadly 'stressful', but the vigilance task is also fatigue-inducing.

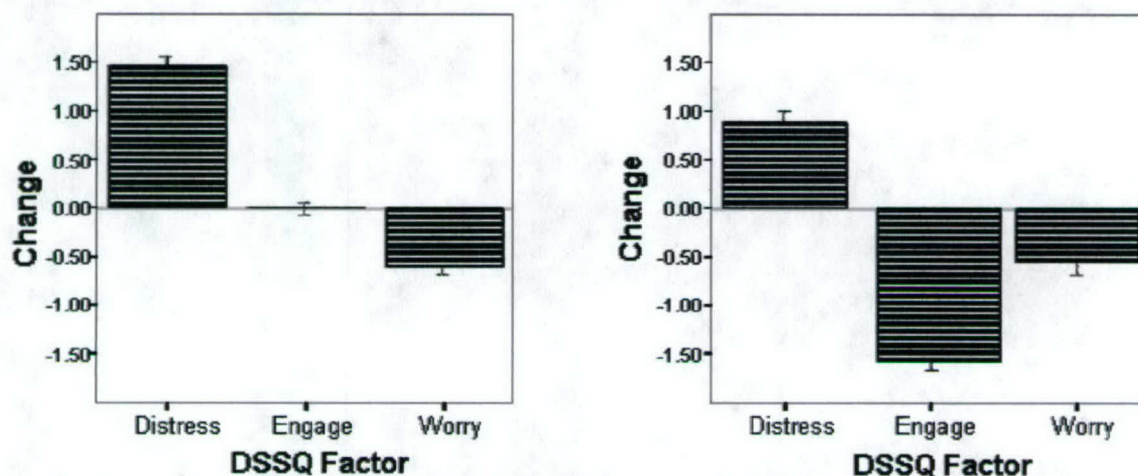


Figure 11. Task-induced change in three DSSQ factors following performance of short task battery (left panel) and vigilance (right panel). Engage. = Task Engagement.

**Workload.** Figure 12 shows workload ratings on the six scales of the modified NASA-TLX (0-10 scales). Ratings for the short battery (left panel) confirm that the tasks are rated as demanding in every respect except physical demand, with the highest ratings assigned to mental and temporal demand. Ratings for the vigilance task (right panel), which are also high, are similar to those seen for other demanding vigilance tasks, with mental demand, effort, and frustration rated as the major contributors to workload.

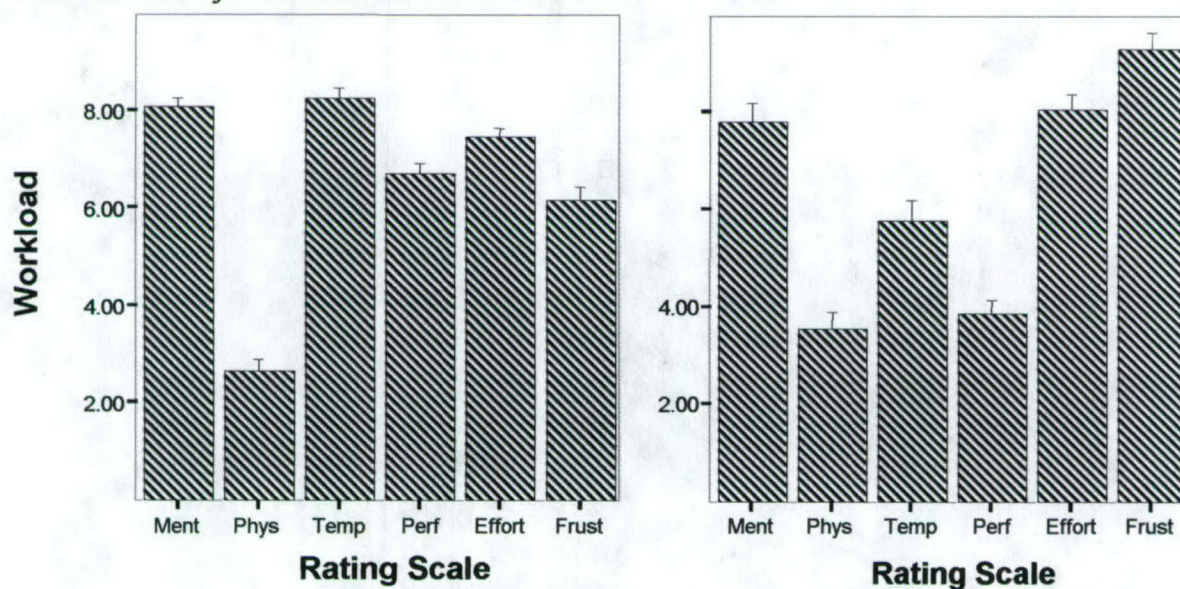


Figure 12. Workload ratings for short battery of tasks (left panel) and the vigilance task (right panel). Scales are Mental Demand (Ment), Physical Demand (Phys), Temporal Demand (Temp), Performance (Perf), Effort, and Frustration (Frust).



*Vigilance performance data.* It is expected that the vigilance task would show a decrement in correct detections, as in previous studies using this task (e.g., Hitchcock et al., 2003). Vigilance performance data are calculated for four successive 10-min periods. Mean percentages of correct signal detections for each time period, shown in Figure 13, suggest a pronounced vigilance decrement. A one-way ANOVA, with task period as a within-subjects factor (4 levels), confirmed a significant effect of task period ( $F(2,226) = 22.52, p < .01$ ). In this and all subsequent ANOVAs, Box's epsilon is used when appropriate in calculating degrees of freedom for repeated measures factors to correct for violations of the sphericity assumption (Maxwell & Delaney, 2004). Other data show that false alarms rates are low in the vigilance task, increasing from 0.56% (first period) to 1.35% (fourth period). Response times tended to increase during the period of watch from a mean of 287 ms (first period) to 471 ms (fourth period). Thus, the temporal decline in detections appears to reflect loss of perceptual sensitivity rather than a change in response criterion (which would depress false alarm rate) or a change in speed-accuracy tradeoff (which would influence response time).

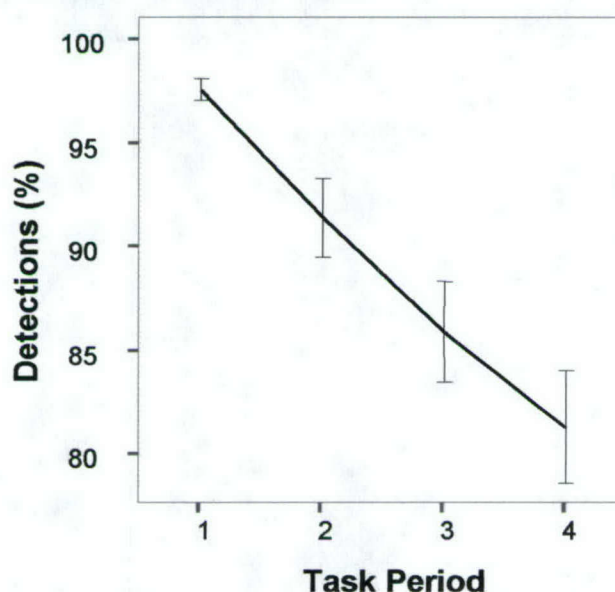


Figure 13. Correct detections (%) as a function of four 10-minute task periods.

#### *Individual differences in blood flow*

The blood flow data raise several key questions:

- Are there reliable individual differences in the different blood flow indices taken during the experiment (i.e., baseline, phasic response to short tasks, blood flow during vigilance)?
- How are different blood flow indices correlated (e.g., how are right vs. left hemisphere measures correlated, baseline vs. phasic response)?
- Do the different short tasks show differential phasic response?
- Does blood flow decline during the vigil?

We will address these in turn, following a brief account of the raw data. It should be noted that due to missing data, N's for these analyses vary from 69 – 92. Due to the length of the study, the TCD signal was lost during recording, on one or both sides, for some observers.



*Initial analyses: Raw data.* Figure 14 shows raw blood flow means in cm/sec for the initial blood flow indices, with 2 consecutive minutes shown for each index. The sequence of indexes from left to right corresponds to the temporal sequence in which measures were taken. BSL-1 refers to the last two minutes of initial baseline recording, which was followed by the first short task (TK-1), and two further episodes of baseline recording (BSL-2, BSL-3), followed by short task performance (TK-2, TK-3). Note that the three tasks – line length discrimination, working memory, and tracking – were presented in counterbalanced order, so that the figure does not differentiate between different tasks. The first two minutes of blood flow recording during vigilance are shown at the right of the figure. Left hemisphere measures are shown by the solid lines; right hemisphere measures by the broken lines.

The figure shows that in the data collected thus far blood flow during short task performance is consistently higher than during the baseline episodes of recording, in which participants have no task imperative. Further, within each 2-min period of recording, means appear to be fairly stable from minute 1 to minute 2. Right hemisphere blood flow tends to be a little lower than left hemisphere blood flow. Levels of blood flow during the first two minutes of the vigilance task are relatively low, similar to baseline levels. The magnitudes of the standard error bars indicate considerable individual variation in blood flow across individuals.

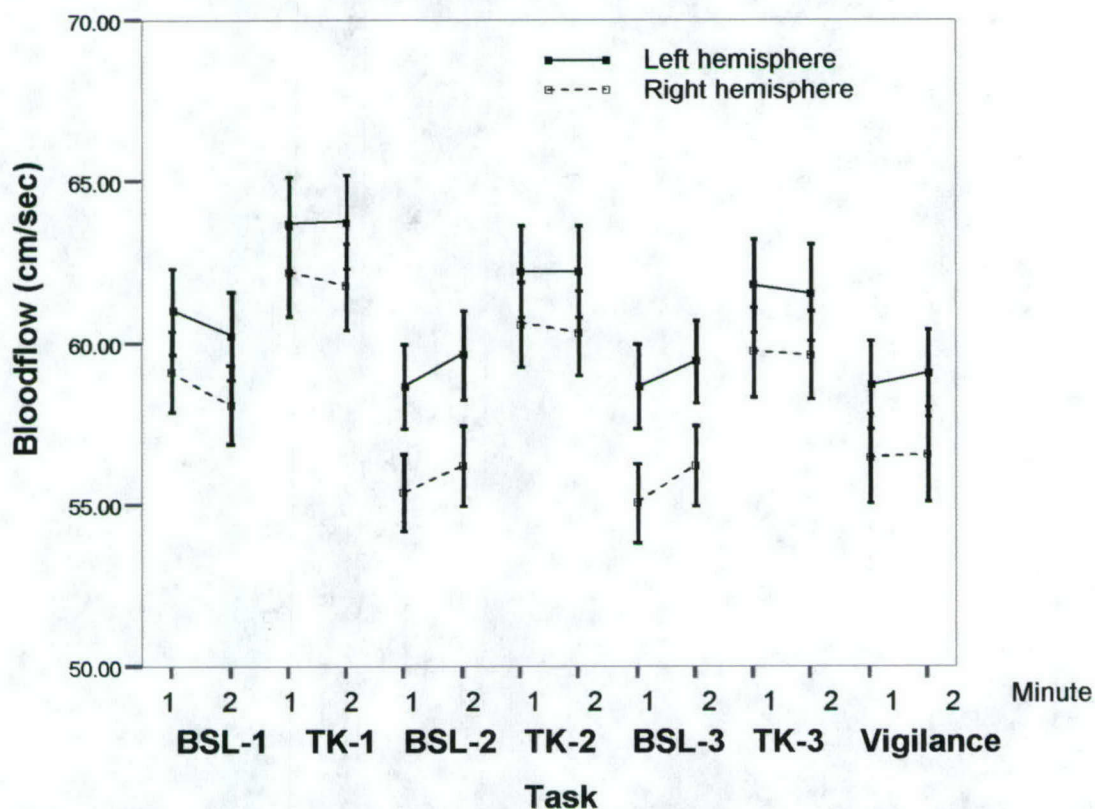


Figure 14. Raw blood flow values (in cm/sec) in several baseline and task conditions, arranged in temporal sequence, as a function of hemisphere. Tasks. BSL-1 = Baseline 1, TK-1 = Task 1, BSL-2 = Baseline 2, TK-2 = Task 2, BSL-3 = Baseline 3, TK-3 = Task 3, Vigilance = Vigilance Task.



### *Individual differences in blood flow: Reliability*

In view of the substantial individual differences in baseline, the next step in the analysis is to convert the raw hemovelocity scores to proportions of the participant's initial resting baseline, assessed as the last two minutes of baseline recording (BSL-1 in Figure 14). Following this step, we intend to calculate the phasic response magnitude produced by each of the three tasks. There are two ways in which this might be done. First, we could find the difference between blood flow during task performance and initial resting baseline, i.e., comparing each task-induced response with BSL-1. Second, we could find the difference between blood flow during the task and the baseline period immediately preceding it, i.e., TK-1 - BSL1, TK2 - BSL2, etc. (Note that the tasks are performed in counterbalanced order, so that TK-1 might refer to any of the three tasks, depending on ordering).

One means for determining which baseline is to be preferred is to check the intercorrelations of the phasic response measures. To the extent that blood flow response is driven by workload, rather than by specific task characteristics, it is expected that response magnitudes will be correlated, at least within the same hemisphere. Two separate calculations of response magnitudes are performed: (1) with reference to the immediately preceding baseline period, and (2) with reference to the initial baseline. In each case, the difference between mean task blood flow and mean baseline blood flow is expressed as a proportion of the participant's relevant baseline value. Each response index is calculated for each of the three tasks: line length detection, working memory and tracking.

Correlations between the task-induced response measures are given in Table 8. Correlations *above* the leading diagonal refer to responses calculated with reference to the initial baseline; correlations *below* the leading diagonal refer to responses calculated with reference to the immediately preceding baseline. For both hemispheres, intercorrelations are higher using the former measure: correlations are typically about .5 - .6 for the 'initial baseline' indices, and about .3 for the 'preceding baseline' indices. As a further check, we also calculate alpha coefficients for each set of three response indices. Values are 0.76 (left) and 0.83 (right) for the 'initial baseline' indices, and 0.57 (left) and 0.55 (right) for the 'preceding baseline' indices. In other words, calculating phasic response using the initial baseline provides a more coherent set of indices, that may measure the individual's responsiveness to various forms of workload. These indices are employed for the remainder of the data analysis.

The preceding analyses suggest that task-induced blood flow response may be reliably assessed. We also check the other blood flow indices for reliability. Initial baseline measures taken from the fourth and fifth minutes of recording (indicated by minutes 1 and 2 in Figure 14 above) correlated at .98 (left) and .97 (right), indicating that individual differences in baseline are highly stable. To check the reliability of blood flow measured during vigilance, mean blood flow, as a proportion of baseline, is calculated for each of four successive 10-min periods of vigilance task performance. The range of correlations is .79 - .92 (left) and .72 - .93 (right), indicating that individual differences in blood flow are highly consistent.



Table 8. Intercorrelations of phasic blood flow indices for each hemisphere, calculated using two different baselines (see text for details). Tasks: Lines = line length discrimination, WM = working memory, Track = tracking.

		Left Hemisphere			Right Hemisphere		
		Lines-P	WM-P	Track-P	Lines-P	WM-P	Track-P
Lines-P	<i>r</i>	---	.419**	.622**	---	.590**	.644**
	N		86	84		83	79
WM-P	<i>r</i>	.290**	---	.597**	.112	---	.658**
	N	87		89	84		84
Tracking-P	<i>r</i>	.334**	.320**	---	.454**	.322**	---
	N	84	89		79	84	

Note. \*  $p < .05$ , \*\*  $p < .01$

*Intercorrelations of blood flow indices.* There are two principal questions concerning intercorrelations of the different blood flow indices. First, are individual differences in the left and right hemisphere measures related? Second, are baseline measures related to measures obtained during task performance? Table 9 shows the inter-hemispheric correlations. Although baseline measures are quite highly intercorrelated, the performance-related indices show reduced, though still substantial, correlations, implying that some part of the variance is hemisphere specific.

Table 9. Correlations between left- and right-hemisphere blood flow indices for five blood flow indices. Baseline = initial baseline blood flow. Lines-P, WM-P and Track-P = phasic responses to line length discrimination, working memory and tracking tasks, respectively. Vigil (1) and Vigil (4) = mean blood flow during periods 1 and 4 of vigilance performance, respectively.

	Baseline	Lines-P	WM-P	Track-P	Vigil (1)	Vigil (4)
<i>r</i>	.741**	.353**	.545**	.603**	.491**	.441**
N	88	80	87	83	86	86

Note. \*  $p < .05$ , \*\*  $p < .01$

Table 10 shows the correlations between selected blood flow measures, separately for left hemisphere indices (above the diagonal) and right hemisphere indices (below the diagonal). It will be seen that the five task-related indices are only weakly correlated with the initial baseline measure, but are substantially intercorrelated among themselves. All three phasic blood flow indices correlate with blood flow during vigilance, with little evidence for differential prediction across the three different tasks.



Table 10. Intercorrelations of selected blood flow indices for left (above leading diagonal) and right (below leading diagonal) hemispheres. Baseline = initial baseline blood flow. Lines-P, WM-P and Track-P = phasic responses to line length discrimination, working memory and tracking tasks, respectively. Vigil (1) and Vigil (4) = mean blood flow during periods 1 and 4 of vigilance performance, respectively.

		Baseline	Lines-P	WM-P	Track-P	Vigil (1)	Vigil (4)
Baseline	<i>r</i>		-.032	-.275**	-.135	-.261*	-.088
	N		86	92	89	86	86
Lines-P	<i>r</i>	.032		.419**	.622**	.389**	.375**
	N	83		86	84	81	81
WM-P	<i>r</i>	-.235*	.590**		.597**	.560**	.464**
	N	89	83		89	86	86
Tracking-P	<i>r</i>	-.137	.644**	.658**		.501**	.409**
	N	85	79	84		84	84
Vigilance (1)	<i>r</i>	-.092	.665**	.676**	.725**		.787**
	N	72	68	72	70		86
Vigilance (4)	<i>r</i>	-.069	.400**	.503**	.421**	.750**	
	N	72	68	72	70	72	

Note. \*  $p < .05$ , \*\*  $p < .01$

In general, the correlational data suggest (1) strong dissociation between baseline and task-related blood flow indices, (2) moderate dissociation between left- and right-hemisphere indices, and (3) little dissociation between different task indices as predictors of blood flow during vigilance.

*Effects of task type and hemisphere on phasic blood flow response.* Figure 15 shows the effects of blood flow and hemisphere on the task-induced phasic response, expressed as a percentage of initial baseline. The data suggest that in both hemispheres the lines task produces a stronger response than tracking. In contrast, the effect of working memory is lateralized, eliciting a stronger left than right hemisphere response. These impressions are confirmed by a 3 x 2 (task x hemisphere) ANOVA, with repeated measures on both factors, and using Box's correction. The main effect of task is significant ( $F(2, 137) = 8.53, p < .01$ ), as is the task x hemisphere interaction ( $F(2, 134) = 15.95, p < .01$ ). Given its working memory component, it is not surprising that the blood flow response shows left lateralization in the working memory task. The other two nonverbal tasks both show a trend towards right lateralization, although a stronger effect of hemisphere might have been expected.



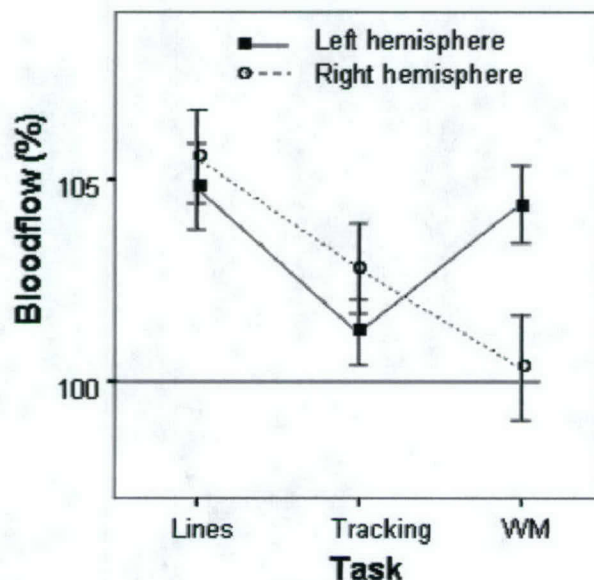


Figure 15. Phasic blood flow response (% baseline) as a function of task type and hemisphere. Lines = line length discrimination, WM = working memory.

*Effects of task period on blood flow during vigilance.* Figure 16 shows per cent blood flow relative to initial baseline as a function of task period (10-minute) and hemisphere.

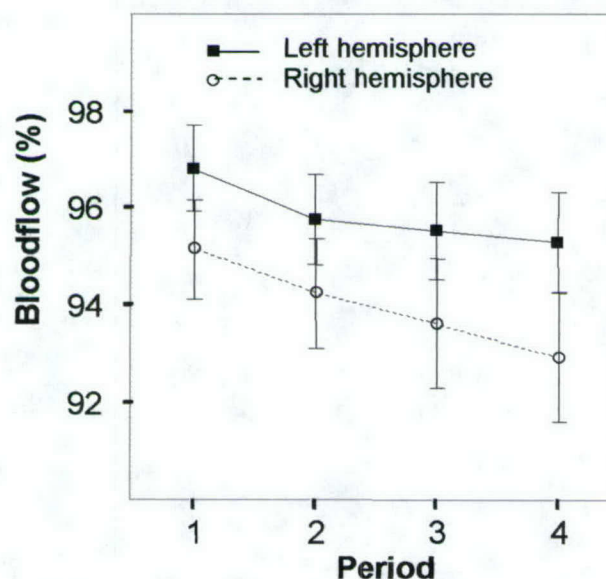


Figure 16. Blood flow (% initial baseline) during task performance as a function of 10-minute task period and hemisphere.

A 2 x 4 (hemisphere x period) repeated measures ANOVA shows a main effect of period ( $F(2, 129) = 5.19, p < .01$ ), consistent with the temporal decline seen in both hemispheres in Figure 16. The tendency for blood flow in Figure 16 to be generally lower in the right hemisphere approaches significance ( $p = 0.065$ ).



### Predictors of vigilance performance

*Correlations between subjective measures and blood flow.* Table 11 displays correlations between the DSSQ factors and the phasic blood flow response indices (correlations between the DSSQ and baseline blood flow are close to zero). Subjective task engagement is predictive of the blood flow response to task performance. Observers high in task engagement on the initial baseline measure tended to show larger blood flow responses on four of six measures. Baseline engagement appears to be especially predictive of right hemisphere response. A similar, but slightly weaker pattern of correlations is found for task-induced engagement, i.e., engagement measured immediately following performance of the short battery of tasks.

Table 11. Correlations between DSSQ factors, measured at baseline, and following performance of the short task battery, with six phasic blood flow response measures. Lines = line length discrimination, WM = working memory, Track = tracking, -L = left hemisphere, -R = right hemisphere.

	DSSQ factor		Lines-L	WM-L	Track-L	Lines-R	WM-R	Track -R
Baseline State	Engagement	<i>r</i>	.100	.190	.274**	.239*	.286**	.309**
		N	86	92	89	83	89	85
	Distress	<i>r</i>	.049	-.073	-.039	-.066	-.173	-.139
		N	86	92	89	83	89	85
	Worry	<i>r</i>	.083	.025	-.089	-.081	.060	.046
		N	86	92	89	83	89	85
Post-task State	Engagement	<i>r</i>	.089	.130	.257*	.268*	.198	.250*
		N	86	92	89	83	89	85
	Distress	<i>r</i>	-.010	-.046	-.159	-.025	-.189	-.145
		N	86	92	89	83	89	85
	Worry	<i>r</i>	.012	.031	-.081	-.144	.022	-.124
		N	86	92	89	83	89	85

Note. \*  $p < .05$ , \*\*  $p < .01$

Correlations between the DSSQ factors and blood flow during vigilance (mean blood flow for each 10-minute period) are also examined. Neither subjective state measured prior to vigilance performance nor post-task state correlate significantly with any blood flow measure,  $p > .05$  in all cases. Individual differences in subjective state appear to relate more to elevations of blood flow, as seen on the short tasks, than to declining blood flow levels, as seen during vigilance.

Correlations of situational awareness (SA: SART) and workload (NASA-TLX) assessed following performance of the vigilance task are also examined. These variables are correlated with post-vigil DSSQ scores, as shown in Table 12. Higher post-task distress is related to lower SA (supply) and to higher overall workload, and worry is positively correlated with demand. However, the correlations between blood flow assessed during vigilance and either SA or workload are not significant,  $p > .05$  in all cases.



Table 12. Correlations between DSSQ factors, measured post-vigil, and situation awareness and workload scales.

DSSQ Factor		SA-Demand	SA-Supply	SA-Complexity	Workload
Engagement	<i>r</i>	.105	.157	.131	-.087
	N	96	97	96	96
Distress	<i>R</i>	.162	-.370**	-.160	.451**
	N	96	97	96	96
Worry	<i>r</i>	.255*	.011	-.072	.121
	N	96	97	96	96

Note. \*  $p < .05$ , \*\*  $p < .01$

*Correlations between subjective measures and performance.* Table 13 displays the correlations between post-vigil DSSQ measures (representing subjective state during vigilance) and detection rate at each of the four 10-minute task periods. These data confirm that performance of vigilance tasks is systematically related to subjective state factors, although, by contrast with previous data, it appears that low distress is a more robust predictor of detection rate than is high engagement. Correlations are also computed for associations between workload and SA and detection measures, but none of these correlations attain significance.

Table 13. Correlations between post-vigil DSSQ factors and detection rates during four 10-minute task periods.

DSSQ Factor		Detections-1	Detections-2	Detections-3	Detections-4
Engagement	<i>r</i>	-.064	-.209	.163	.161
	N	96	96	96	96
Distress	<i>r</i>	-.129	-.287**	-.381**	-.318**
	N	96	96	96	96
Worry	<i>r</i>	-.112	.134	.137	.130
	N	96	96	96	96

Note. \*  $p < .05$ , \*\*  $p < .01$

*Correlations between blood flow and performance.* We have also correlated detection rates on the vigilance task with blood flow assessed at baseline, and with phasic blood flow response indices, and with blood flow assessed during vigilance. These correlations show no consistent trend, and none are statistically significant,  $p > .05$  in all cases. The lack of any functional relation between cerebral blood flow and vigilance performance is surprising. Possibly, functional significance of blood flow varies with the operator's state. Blood flow might be more positively related to performance in operators who are motivated and actively engaged with the task than in those whose interest is low. To explore that possibility, participants were divided arbitrarily into those  $> 1$  SD below the normative mean on post-vigil DSSQ engagement (range: -1.04 - -3.29), and those whose engagement score  $\leq -1$  SD (range: -1.00 - +1.64). The former constitute a group of fatigued, demotivated individuals, while the latter retain at least some interest in the task. Blood flow and vigilance performance are then correlated separately for each group.

Table 14 displays the correlations for each group. Each entry shows the concurrent correlation at a given time period, e.g., the first entry correlates left-hemisphere blood flow during period 1 with detection rate in the same period. Especially for right-hemisphere measures, the two



groups show opposite blood flow performance trends. In the "more disengaged group," blood flow tends to be negatively associated with detection rate, whereas in the "less disengaged group," blood flow performance associations tend to be positive.

Table 14. Correlations between blood flow, measured during vigilance, and detection rates during four 10-minute task periods, in more and less disengaged participants.

Hemisphere		Detections-1	Detections-2	Detections-3	Detections-4
More disengaged participants					
Left	<i>r</i>	-.064	-.209	-.215	-.277
	N	49	49	49	49
Right	<i>r</i>	-.216	-.372**	-.215	-.308
	N	39	39	39	39
Less disengaged participants					
Left	<i>r</i>	.130	-.001	.165	.230
	N	36	36	36	36
Right	<i>r</i>	.304	.236	.593**	.136
	N	32	32	32	32

Note. \*  $p < .05$ , \*\*  $p < .01$

Table 15 displays equivalent correlations using the phasic blood flow response indices for each of the short tasks as predictors for vigilance. Data are for the *right hemisphere only*. (Correlations for left hemisphere indices did not show consistent trends). In the "less disengaged participants," the data show a consistent trend for blood flow response to the short tasks to predict higher detection rates in vigilance. By contrast, in the "more disengaged participants" the trend is towards small negative correlations between blood flow and detection rate. Higher blood flow may only be diagnostic of better performance in individuals who are actively engaged with the task.



Table 15. Correlations between right-hemisphere phasic blood flow response measures and detection rates during four 10-minute task periods, in more and less disengaged participants. Lines = line length discrimination, WM = working memory.

Phasic response		Detections-1	Detections-2	Detections-3	Detections-4
More disengaged participants					
Lines-R	<i>r</i>	-.224	-.193	-.197	-.203
	N	47	47	47	47
WM-R	<i>r</i>	-.027	-.136	-.224	-.122
	N	51	51	51	51
Tracking-R	<i>r</i>	-.174	-.157	-.109	-.081
	N	47	47	47	47
Less disengaged participants					
Lines-R	<i>r</i>	.487**	.408*	.565**	.567**
	N	35	35	35	35
WM-R	<i>r</i>	.364*	.232	.475**	.320
	N	37	37	37	37
Tracking-R	<i>r</i>	.469**	.158	.342*	.199
	N	37	37	37	37

Note. \*  $p < .05$ , \*\*  $p < .01$

It is emphasized that these results are suggestive only. A moderator effect of task engagement was not predicted initially. Should these results remain significant at the end of the study, replication will be required. However, if the diagnosticity of blood flow should prove to depend on subjective motivation and engagement, this result will have important implications for its practical utility.

### Discussion

As this study is ongoing, only a brief discussion of results is presented. The following findings appear to be the most salient:

- Manipulation checks indicated that the tasks were producing the psychological effects anticipated. Both the short battery of tasks and the vigilance task elicited high levels of subjective workload and distress, together with high task disengagement in the case of vigilance. Further, consistent with past research, the vigilance task showed temporal declines in signal detection and in cerebral blood flow.
- The data also suggest that individual differences in blood flow parameters may be reliably assessed. Phasic responses to the short tasks may be assessed separately from baseline blood flow, and phasic response is predictive of blood flow during the subsequent, more prolonged vigilance task.



- It may be important to distinguish left and right hemisphere blood flow indices. Phasic responses appear to show some lateralization, depending on the nature of the task, and equivalent left- and right-hemisphere indices appear to be only moderately intercorrelated.
- Correlations between blood flow, subjective state, and performance tend to be small in magnitude. However, baseline task engagement predicts subsequent phasic blood flow responses, and high distress appears to signal concurrent loss of vigilance. Contrary to one of our main predictions, individual differences in blood flow and detection rate do not appear to be substantially correlated.
- A post hoc analysis suggested that correlations between blood flow and detections may vary qualitatively with the participants' level of task engagement. Right-hemisphere blood flow may be indicative of better performance in more engaged participants, but poorer performance in those that have largely abandoned active efforts to maintain successful performance.



## STUDIES AT GEORGIA STATE UNIVERSITY

We tested over 150 undergraduate participants in pilot research. These studies were designed to validate our "shoot/don't-shoot" ("Watchkeeper") task and to determine the parameters of testing for other cognitive assessments, so that we will be maximally positioned to collect the best data once the TCD is working properly. One study is ongoing, and involves comparison of response-apparatus options for the Watchkeeper task. The other studies were designed to define the optimal Assessment Software for Attention Profiles (ASAP) battery and the relation of this battery to other measures.

### *Watchkeeper study*

As described in the initial proposal, a new criterion task was developed for this research. The Watchkeeper task was inspired by the demands faced by sentries or other individuals who must maintain vigilance on an operational scene, detect the presence of persons who may or may not be threats, and shoot accurately when threats are presented. The Watchkeeper task was modeled after the marksmanship range at the Aberdeen Proving Grounds. A grassy field, complete with trees and hills, is projected as the stimulus background in front of participants, as shown in Figure 17. The display is projected on a screen, and occupies approximately 45 degrees of visual angle (approximately 2.3 meters diagonal). The participant must monitor this scene for 20 minutes, searching for threat images (blue rectangles) that appear infrequently from behind the trees or over the hills and then disappear again behind the blinds. Nonthreat images (orange rectangles) appear more frequently, but with comparable stimulus locations and display durations. The targets are scaled so that distance and size covary as they would in a natural display. Participants see threat or nonthreat stimuli with an event rate averaging 10/minute. The ratio of nonthreat:threat stimuli is 80%:20%.

To respond to these images, participants use one of three manipulanda (assigned as a between-subjects manipulation). Some participants use the computer mouse to click on the threat images when they appear. Others respond with a replica of a 9-mm handgun, modified by LaserShot, Inc. to produce a laser blip on the screen when fired. This laser spot is interpreted by our software as a mouse click, such that the time and location of each shot is recorded. The third group of participants is tested with a rifle as the response manipulandum. The rifle is similarly modified so as to simulate a mouse click on the screen. These response devices are displayed in Figure 17. No response is required to the nonthreat images.

The purpose for this ongoing study is to validate the parameters of the Watchkeeper task and to see whether the effects of time-on-task differ as a function of response apparatus. Our focus is twofold. First, we are concerned that the novelty and game-like character of the weapons might engage participants to an extent that the vigilance decrement might not be observed in the 20-minute vigil. Such findings would necessitate a switch to a response mode with less verisimilitude to the applied context, the utilization of more experienced participants so that novelty effects would be removed, or the employment of a longer vigil.





Figure 17. The Watchkeeper task.

Second, we are concerned that, relative to the mouse, the novelty of the “aim-and-shoot” demands of the pistol and rifle would produce practice effects across the vigil with respect to marksmanship accuracy. That is, participants may actually get better at aiming and firing the weapon across the vigil, confounding any effect from inattention or the stress generated by the sustained attention demands.

As illustrated in Table 16, the *threat-detection rate* remains high across the 20-minute vigil, but response time increases significantly and becomes more variable across the watch. This pattern of results is as predicted in the original proposal. It provides the disruptions in performance that we will examine to determine whether they can be predicted by cognitive profiles and psychophysiological indicators such as TCD. However, consistent with the second concern described above, the *marksmanship data* (i.e., shot or mouse-click accuracy) show a performance improvement in the initial minutes of the task, and then to decay with time-on-task. If this pattern of results holds up in final analyses, we will administer an “aiming practice” task before the Watchkeeper vigil begins. In this way we should be able to eliminate any small but reliable improvement in performance over time.

Table 16. (Preliminary analysis of ongoing pilot research). Watchkeeper task performance for each dependent measure as a function of watch-period.

	Minute 1-5	Minute 6-10	Minute 11-15	Minute 16-20
Mean Hit Rate (Percent correct)	99%	98%	99%	97%
Mean Response time in sec (SD)	1.53 (.33)	1.72 (.33)	1.88 (.35)	1.92 (.47)
Mean Shot Accuracy	335	220	235	268

We have not analyzed these data yet to determine whether the findings differ by group (response manipulandum). However, we remain confident, given this pattern of results, that the Watchkeeper task is a suitable criterion task against which the predictions of TCD and other measures can be assessed.



### *Attention profiles*

In order to accommodate the collection of predictor data in the same experimental session as the Watchkeeper task, it was necessary to streamline the battery of cognitive measures that will be administered. Two pilot studies have been conducted to determine the optimal array of tasks and parameters that make-up the Assessment Software for Attention Profiles (ASAP) battery to be used in the proposed research. We have shown that the performance effects associated with time-on-task may be mediated by the cognitive skills that participants bring to the test session. In particular, we have profiled individual differences in attention skills that characterize participants. At least three component skills constitute the construct of attention: "focusing," "scanning," and "sustaining" (Washburn, Putney, & Tirre, 1998). Focusing refers to the ability to concentrate, to invest mental energy, and to inhibit prepotent responses. This factor of attention is often also called "executive attention." It contrasts with attention scanning (or orienting), which is the ability to move attention in space. This factor is frequently associated with operations like searching or shifting attention. The third factor of attention is sustaining or alerting, and is tapped by tasks that require vigilance or attending over time. The ASAP battery provides measures of ability along each of these dimensions, plus some additional measures associated with executive functioning.

These measures of cognitive ability are important because the effects of combat stress, time-on-task, or other Army-relevant variables may be different for participants with different cognitive profiles. In the first pilot study on the ASAP battery, we demonstrated this difference by showing that participants with the best sustained-attention skills (as measured by a six-minute continuous performance task) learn more rapidly and to a higher asymptote than do participants within the lowest quartile of sustained-attention skills. The training task used here was a threat-detection task modeled after a security-screener's duties. Participants were trained in a 25-minute session to recognize threat images. Each participant also completed a 30-minute version of the ASAP battery. The primary purpose for this study was to validate that the ASAP battery used here, which was shortened from the standard version by reducing trials/task and, in some cases, by removing tasks altogether, would nonetheless produce profiles consistent with those obtained in previous studies. The secondary purpose for this study was to determine whether the effects of practice on the task would be differentially manifest for different groups of participants, where the groups were formed on the basis of cognitive (and specifically, sustained-attention) profiles.



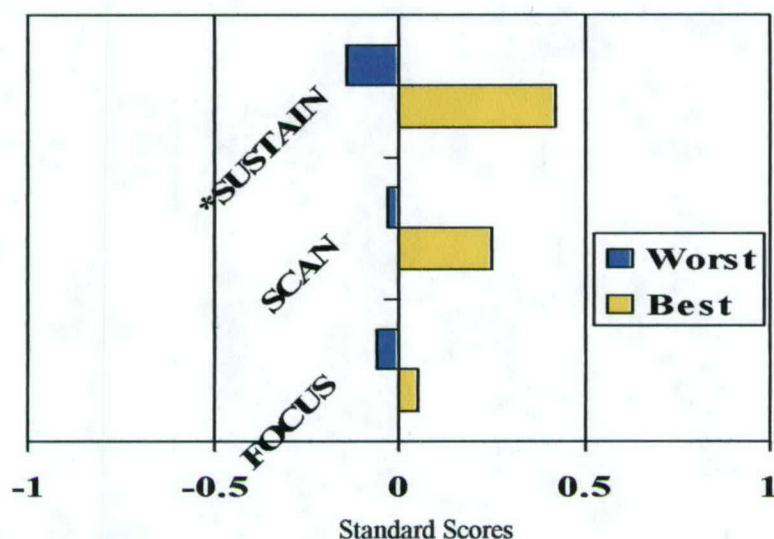


Figure 18. Cognitive Profiles of participants who perform best or worst (i.e., who show the most vs. the least improvement) in threat detection across a 25-minute training session. The profiles are averaged standardized scores for the groups across the ASAP factors of attention.

The results of this pilot study are illustrated in Figure 18. Factor analysis of the shortened ASAP battery did reveal the three latent variables predicted on the basis of previous research. However, the variance accounted for by these three factors was smaller than had been expected. Note that these ability profiles were then successful in predicting performance on the threat-recognition task (more specifically, in predicting how threat-recognition performance would improve with practice).

Parameters were adjusted on the ASAP tasks and a second pilot experiment was conducted. Participants were given the revised ASAP battery and a psychophysical judgment task (speeded numerical judgments) under either low workload conditions or high workload conditions. The pacing of stimulus presentations was used as the manipulation of workload. Although analysis of these data is preliminary, the following tentative conclusions can be offered. First, relative to the short-version used in the first pilot study, the new parameters of the ASAP battery produce better measures of the attentional skills of the participants. The factor structure revealed in these data is appropriate to predictions, as is the variance accounted for. Second, it again appears that the effects of the manipulation of interest (in this case, cognitive workload) are modulated by the participants' attention skills. As illustrated in Figure 19, participants within the top quartile for *attention-focusing* skills showed smaller effects of the workload manipulation than did participants in the bottom quartile of attention-focusing. Similar analyses will be conducted for the other two factors of attention ability. Additionally, we will use multivariate analyses (regression and path analysis) to determine the predictive and moderating relation between individual differences in attention performance and the effects of workload.



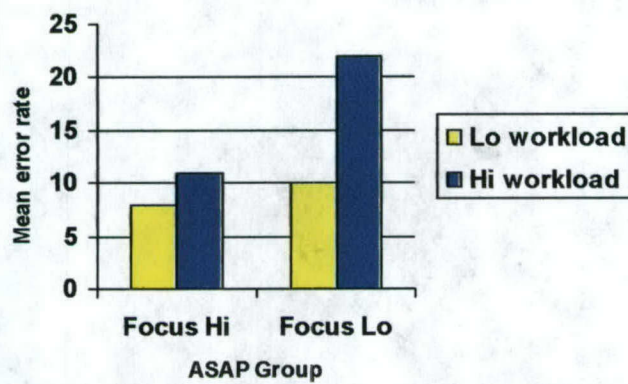


Figure 19. Errors on a psychophysical judgment task, as a function of level of workload and participants' attention-focusing skills.



## IMPLICATIONS FOR FUTURE RESEARCH AND TIMELINE

The overall strategy for our research is to conduct several substantial studies comparing multiple predictors of sustained performance (e.g., Study 3), as set out in our original proposal, supplemented by additional studies exploring specific aspects of tasks and predictors in more detail (e.g., Studies 1 and 2).

*University of Cincinnati.* At this point, research at the University of Cincinnati (UC) is a little behind the timeline set out in the original proposal, in that we had hoped to finish our first substantial study of predictive validity during the first year.

There are two principal factors contributing to the delay. First, we did not receive approval from the Army Human Subjects Research Review Board until February 13<sup>th</sup>, 2004. Consequently, we could not accomplish much testing in the first academic quarter of the year. As our participants are introductory psychology students, our testing is time-locked to the academic calendar. Second, the experimental protocol is lengthier and more complex than those used in previous studies of cerebral blood flow, owing to the need to assess multiple predictors. In addition, participants have to meet stringent exclusion criteria, including avoiding certain foods and beverages, to meet requirements for cortisol assays. For these reasons, we have had more participants fail to attend or complete test sessions than we originally anticipated. However, we anticipate concluding the ongoing study around February 2005.

We will review our experiences with the current protocol before designing the next study in the sequence. It may be possible to simplify some aspects, such as using shorter performance task. We will also evaluate the utility of measuring cortisol as a potential predictor of performance, when these results are received. If cortisol is not predictive, the protocol and exclusion criteria may be simplified accordingly. We will certainly complete one further study comparing multiple predictors at UC in the second year of the study. This study will investigate predictors of a cognitive vigilance task, imposing a substantial working memory task, as a complement to the sensory vigilance task used in the current research (as described in the original proposal). We are hopeful also that we will begin a study investigating cerebral blood flow as a predictor of sustained driving performance, using the simulator methods described in Study 2. This timeline is contingent upon receiving reasonably rapid human subjects approvals from local and Army review boards.

*Georgia State University.* It has been a frustrating year for the investigators of the Georgia State University (GSU) portion of the project. Funds were not available at GSU for expenditure until March of 2004, and the transcranial Doppler (TCD) device was not installed at that facility until the end of June, 2004. Thus, we could not run participants until the beginning of Fall, 2004 semester. (Few students are available in the participant pool during the Summer terms.). Even in the Fall, however, we have been unable to begin the proposed experiment, owing to software problems with the TCD unit that have yet to be resolved by the manufacturer. Accordingly, we have spent little of our Year-1 funding. The expenditures include funds associated with the purchase of the TCD and stimulus-presentation hardware; funds to support the graduate assistant and the PI who have conducted the preliminary investigations; and travel funds for the graduate assistant to visit the University of Cincinnati for training with the TCD



equipment. However, we have achieved several essential accomplishments during this short and frustrating period of funding. Laboratory space has been renovated for the project. The team of investigators has learned to use the TCD hardware, and each researcher has practiced the placement and calibration of the sensors. The "shoot/don't-shoot" vigilance task has been developed and piloted. The cognitive-task battery has been tested, refined, and re-tested. As a consequence, the GSU portion of this project is now ready to begin the proposed study as soon as the TCD software problem is resolved. The fact that our design and methods can be informed and revised, where necessary, by the initial results produced by our colleagues in Cincinnati is a great benefit.

*Conclusion.* We are confident that the overall aims of the project, to test potential diagnostic predictors of current and future performance, may be met within the three-year time period originally envisaged. Initial results are already providing us with information on what measures of cerebral blood flow may tell us about operators' fitness to perform a sensory vigilance task. In the next year of the research, we anticipate being able to confirm and extend these findings to additional tasks, including the sustained rifle marksmanship tasks developed at GSU during the past year.



## KEY RESEARCH ACCOMPLISHMENTS

- The results of Study 1 confirm the high demands of detection tasks based on missing stimulus elements. Tasks of this kind – such as military surveillance – may be especially vulnerable to temporal decrements in cerebral blood flow and performance. They also elicit subjective stress.
- The research confirms the utility of subjective state measures as predictors of performance. It has been shown that a state questionnaire (the DSSQ) predicts driver performance in high workload conditions (Study 2), phasic blood flow response to highly demanding tasks (Study 3), and vigilance (Study 3). Results from Study 2 suggested that state measures may be diagnostic of fitness to perform in both stressful and nonstressful environments.
- Study 3 demonstrated the feasibility of measuring individual differences in cerebral blood flow and distinguishing psychometrically between baseline and phasic responses. The data further suggest that blood flow measured during performance reflects psychological as well as physiological factors.
- Tasks have been developed corresponding to military situations including monitoring for the disappearance of a stimulus element, as in surveillance (Study 1), maintaining vehicle control while simultaneously monitoring the environment for hazards (Study 2), and monitoring a screen resembling a radar display (Study 3).
- Findings confirm the utility of resource theory as a theoretical framework for understanding workload effects on vigilance and blood flow (Study 1), and multi-tasking and automation in vehicle driving (Study 2). The utility of the theory as the basis for identifying diagnostic predictors of performance remains uncertain, although the utility of subjective task engagement as a performance indicator may be accommodated by resource theory.
- Preliminary results from Study 3 raise the possibility that the diagnosticity of cerebral blood flow (and perhaps other physiological indices) depend on the operator's task commitment and involvement. If this result is substantiated and generalized, this finding would have important implications for future research on diagnostic indices.
- Pilot data from studies conducted at Georgia State University demonstrate vigilance-like decrements in performance of a fairly realistic rifle marksmanship task. Performance on this task appears to be predictable from measures of basic attentional abilities.



## REPORTABLE OUTCOMES

### *Conference presentations*

Hollander, T., Warm, J.S., Matthews, G., Shockley, K., Dember, W.N., Weiler, E.N., & Scerbo, M. (2004, April). *Feature presence/absence and the event rate effect in sustained attention*. Paper presented at the 96th Annual Meeting of the Southern Society for Philosophy and Psychology, New Orleans, LA.

Hollander, T., Warm, J.S., Matthews, G., Shockley, K., Dember, W.N., Weiler, E.N., Tripp, L.D., & Scerbo, M.W. (2004). Feature presence/absence modifies the event rate effect and cerebral hemovelocity in vigilance performance. *Proceedings of the Human Factors and Ergonomics Society*, 48, 1943-1947.

Matthews, G. (2004, October). *Picking a bone with Spearman: A contrary view of mental energy and attention*. Invited address, 4<sup>th</sup> Spearman Conference on Diagnostics for Education, Philadelphia, PA.

Warm, J.S. (October, 2004). *Cerebral Hemodynamics and Vigilance Performance*. Invited address, National Research Council Committee on Human Factors, Washington, DC.

### *Accepted conference presentations*

Funke, G., Matthews, G., Emo, A.K., & Warm, J.S. (2005, March). *The effects of partial-vehicle automation on driver mood and performance in a simulated winter drive*. Paper accepted for presentation at the 97th Annual Meeting of the Southern Society for Philosophy and Psychology, Durham, NC.



## CONCLUSIONS

Study 1 investigated the status of cerebral blood flow as a concurrent indicator of operator status, in the novel context of a vigilance task requiring detection of the absence of a key stimulus element. Both performance and subjective stress data confirmed the potential for human factors problems arising from tasks requiring monitoring for 'absence' rather than 'presence'. The study confirmed earlier findings of temporal decline in blood flow during performance of sustained signal detection tasks. It added to these results by showing that, in the first 10 minutes of the period of watch, there was a steeper decline in blood flow in the feature-absent condition, relative to the feature-present condition. The study points towards the potential value of diagnostic monitoring of functional status during military operations such as surveillance of groupings of enemy forces.

Study 2 tested the utility of subjective measures as predictors of driving performance in a dual-task situation requiring signal detection in stressful and nonstressful conditions. Level of vehicle automation was also manipulated. Effects of automation on performance could, in general, be accommodated within attentional resource theory. The study did not reveal any special human factors issues arising from the interaction of stress and automation. Analyses of individual differences suggested that, as in previous studies, standard deviation of lateral position was a sensitive index of driver performance. Optimizing prediction of this performance metric required use of multiple predictors including stable personality traits as well as short-duration subjective states. The emergence of subjective task engagement as a predictor links these results to previous studies of vigilance. At a practical level, the findings suggest that both low task engagement and high worry may be sources of impairment during sustained vehicle operation, including military convoy driving.

Study 3 is an ongoing study, and only tentative and preliminary results are offered at this point. It is designed to compare various subjective and physiological measures as diagnostic indices of concurrent and future fitness to perform a demanding sustained attention task, using a display resembling those used in air traffic control. Data collected so far suggest three conclusions. First, manipulation checks suggest that task paradigms are functioning as intended in provoking high levels of stress and workload. The vigilance task shows significant temporal decrements in both detection rate and cerebral blood flow. Second, reliable measurement of individual differences in blood flow has been demonstrated; measures associated with task performance are quite distinct from baseline measures obtained in the absence of a task imperative. Third, some meaningful correlations between subjective state and both blood flow and performance have been obtained. However, blood flow indices do not appear to correlate with vigilance performance. Further analyses suggest that the diagnosticity of blood flow indices may vary with the individual's task engagement. Elevated blood flow may only predict vigilance to the extent that the person remains engaged and committed to performing well.

The pilot studies conducted at Georgia State University have been primarily methodological in character. They have focused on the development and validation of tasks and procedures for use in subsequent, larger scale studies to be initiated early in 2005. This work has produced a rifle marksmanship simulation, the Watchkeeper task, that requires participants to



evaluate threat and to make 'shoot/don't shoot' decisions. It has also refined an existing battery of short performance tests (the ASAP battery) used for evaluating attentional abilities within a validated multidimensional model, to be used in future TCD studies. Data from the pilot studies also suggest two important, substantive findings. First, the Watchkeeper task shows performance deficits over time that may be analogous to the vigilance decrement, and so may relate to cerebral blood flow. Second, performance on the task has been shown to relate to at least one of the dimensions of attentional ability measured by ASAP, attentional-focus.

*Scientific impact.* The theory that guides much previous work in this field, including our own previous work on cerebral blood flow, is attentional resource theory. The theory proposes that mental workload corresponds directly to the quantity of attentional resources invested in information-processing and performance. It follows that, in both laboratory and applied settings, the likelihood of performance impairment and breakdown may be gauged from an assessment of the total of demands on resources, including immediate task demands and depletion of resources due to fatigue and stress. The promise of transcranial Doppler sonography and other diagnostic techniques (both physiological and subjective) is that they may indicate the extent to which the individual's available resources are adequate for handling task demands.

The results so far indicate both strengths and weaknesses of this resource/workload approach to predicting performance impairment. Data from all three studies broadly support the utility of this theory for predicting effects of task demands on performance, including driver performance (Study 2) and vigilance (Studies 1 and 3). Temporal decline in blood flow (Studies 1 and 3) may also be understood within a resource theory framework. However, measuring individual differences in resource availability is challenging. Although subjective measures, including subjective task engagement (Study 2), continue to show promise, data obtained so far from Study 3 do not show robust blood flow – performance associations. Factors such as task commitment that may moderate the diagnostic significance of cerebral blood flow may merit further investigation.

The research has also accomplished several methodological advances. In terms of performance assessment, we have confirmed the validity of a 'feature-absent' vigilance task as a tool for investigating the vigilance decrement (Study 1). Study 2 established techniques for investigating multi-tasking during driving, combining a tracking (psychomotor control) component with hazard monitoring in a simulated driving task. The study also validated methods for manipulating stress and level of automation in driving. For the first time, we have shown (Study 3) that individual differences in multiple blood flow indices may be reliably assessed and differentiated from one another, extending the range of potentially diagnostic psychophysiological tools. It has also been demonstrated that simultaneous bilateral recording of cerebral blood flow is feasible (at the cost of a slightly increased likelihood of signal loss during prolonged recording).

Finally, although this research is basic in nature, it may support future applied work. Practical issues to which the research is relevant include military surveillance (Study 1), vehicle operation (Study 2) and diagnosis of the soldier's fitness to perform sustained monitoring tasks, such as sentry duty and radar monitoring. Future studies will be necessary to identify optimal predictor sets for specific tasks – and, as suggested by first results from Study 3 – moderator



factors that control the diagnostic validity of specific indicators. However, results so far generally support the strategy of using multiple indicators to maximize predictive validity.



## References

- Aaslid, R. (1986). Transcranial Doppler examination techniques. In R. Aaslid (Ed.), *Transcranial Doppler sonography* (pp. 39-59). New York: Springer-Verlag.
- Adams, R.J., Nichols, F.T., Hess, D.C. (1992). Normal values and physiological variables. In D.W. Newell & R. Aaslid (eds.), *Transcranial Doppler* (pp.41-66). New York: Raven Press.
- Beam, C., Warm, J.S., Matthews, G., Dember, W.N., Hitchcock, E.M., Tripp, L., & Parasuraman, R. (2002, October). *Effects of vigilance task type on cerebral blood flow velocity and subjective stress*. Paper presented at the 44th Annual Meeting of the Human Factors and Ergonomics Society, Baltimore, MD.
- Bohnen, N., Houx, P., Nicolson, N., & Jolles, J. (1990). Cortisol reactivity and cognitive performance in a continuous mental task paradigm. *Biological Psychology*, 31, 107-116.
- Born, J., Hitzler, V., Pietrowsky, R., Pauschinger, P. (1988). Influences of cortisol on auditory evoked potentials (AEPs) and mood in humans. *Neuropsychobiology*, 20, 145-151.
- Buckingham, J.C., Gillies, G.E., & Cowell, A.M. (1997). *Stress, stress hormones and the immune system*. Chichester: Wiley.
- Caplan, L.R., Brass, L.M., DeWitt, L.D., Adams, R.J., Gomex, C., Otis, S., Weschler, L.R., & von Ruef, G.M. (1990). Transcranial Doppler ultrasound: Present status. *Neurology*, 40, 696-700.
- Coull, J., Frackowiak, R.S.J., & Frith, C. (1998). Monitoring for target objects: activation of right frontal and parietal cortices with increasing time on task. *Neuropsychologia*, 36, 1325-1334.
- Davies, D.R. & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.
- Davies, D.R., & Tune, G.S. (1969). *Human vigilance performance*. New York: Elsevier.
- Dickerson, S.S., Kemeny, M.E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130, 355-391.
- Duschek, S., & Schandry, R. (2003). Functional transcranial Doppler sonography as a tool in psychophysiological research. *Psychophysiology*, 40, 436-454.
- Endsley, M.R. (1995) Towards a theory of situation awareness in dynamic systems. *Human Factors*, 37,85-104.
- Galinsky, T.L., Rosa, R.R., Warm, J.S., & Dember, W.N. (1993). Psychophysical determinants of stress in sustained attention. *Human Factors*, 35, 603-614.



- Gray, J. A. (1987). *The psychology of fear and stress* (2nd ed.). New York: Cambridge University Press.
- Gulian, E., Glendon, A. I., Matthews, G., Davies, D. R., & Debney, L. M. (1990). The stress of driving: A diary study. *Work and Stress*, 4, 7-16.
- Hancock, P.A., & Hart, S.G. (2002). Defeating terrorism: What can human factors/ergonomics offer. *Ergonomics and Design*, 10, 6-16.
- Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, 31, 519-537.
- Harms, L. (1986). Drivers' attentional responses to environmental variations: A dual-task real traffic study. In A.G. Gale, M.H. Freeman, C.M. Haslegrave, P. Smith, & S.P. Taylor (Eds.), *Vision in vehicles* (pp. 131-138). Amsterdam: North-Holland.
- Hart, S.G., & Staveland, L.E. (1988). Development of a multi-dimensional workload scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North-Holland.
- Hellige, J.B. (1993). *Attentional asymmetry: What's right and what's left*. Cambridge, MA: Harvard University Press.
- Helton, W.S., Warm, J.S., Matthews, G., & Corcoran, K.J. (2002). Further tests of an abbreviated vigilance task: Effects of signal salience and jet aircraft noise on performance efficiency and stress. *Proceedings of Human Factors and Ergonomics Society*, 1456-1461.
- Hitchcock, E.M., Dember, W.N., Warm, J.S., Moroney, B.W., & See, J.E. (1999). Effects of cueing and knowledge of results on workload and boredom in sustained attention. *Human Factors*, 41, 365-372.
- Hitchcock, E.M., Warm, J.S., Matthews, G., Dember, W.N., Shear, P.K., Tripp, L.D., Mayleben, D.W., & Parasuraman, R. (2003). Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science*, 4, 89-112.
- Hockey, G.R.J. (1986). A state control theory of adaptation to stress and individual differences in stress management. In G.R.J. Hockey, A.W.K. Gaillard, & M.G.H. Coles (Eds.), *Energetics and human information processing*. Dordrecht: Martinus Nijhoff.
- Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, 45, 73-93.



- Hollander, T.D., Warm, J.S., Matthews, G., Dember, W.N., Parasuraman, R., Hitchcock, E.M., Beam, C.A., & Tripp, L.D. (2003, November). *Effects of signal regularity and salience on vigilance performance and cerebral hemovelocity*. Paper presented at the 44th Annual Meeting of the Psychonomic Society, Vancouver, BC.
- Huwe, S., Hennig, J., & Netter, P. (1998). Biological, emotional, behavioral, and coping reactions to examination stress in high and low state anxious subjects. *Anxiety, Stress & Coping: An International Journal*, 11, 47-65.
- Johnson, A., & Proctor, R.W. (2004). *Attention: Theory and practice*. Thousand Oaks, CA: Sage.
- Kanki, B. G. (1996). Stress and aircrew performance: A team-level perspective. In J. E. Driskell & E. Salas (Eds.), *Stress and human performance* (pp.127-162). Mahwah, NJ: Erlbaum.
- Kantowitz, B. H., & Simsek, O. (2001). Secondary-task measures of driver workload. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 133-163). Mahwah, NJ: Erlbaum.
- Kinomura, S., Larsson, J. Gulyas, B., & Roland, P.E. (1996). Activation by attention of the human reticular formation and thalamic intralaminar nuclei. *Science*, 271, 512-515.
- Kirk, R.E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole
- Kirschbaum, C., & Hellhammer, D.L. (2000) Salivary cortisol. In G. Fink (Ed.), *Encyclopedia of Stress: Vol. 3*. New York: Academic.
- Klingelhöfer, J., Sander, D., & Wittich, I. (1999). Functional ultrasonic imaging. In V.L. Babikin & L.W. Wechsler (Eds.) *Transcranial Doppler Ultrasonography* (pp. 49-66). Boston: Butterworth-Heinemann.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.
- Lundberg, U., & Johansson, G. (2000). Stress and health risks in repetitive work and supervisory monitoring work. In R. W. Backs & W. Boucsein (Eds.), *Engineering Psychophysiology* (pp. 339-360). Mahwah, NJ: Erlbaum.
- Lupien, S., Lecours, A.R., Lussier, I., Schwartz, G., et.al. (1994). Basal cortisol levels and cognitive deficits in human aging. *Journal of Neuroscience*, 14, 2893-2903.
- Mackworth, N.H. (1950/1961). Researches on the measurement of human performance. *Medical Research Council Special Report Series 268*. London: HM Stationary Office (reprinted from *Selected papers in the design and use of control systems*, pp. 174-333, by H.W. Sinaiko, Ed., 1961, New York: Dover).



- Mason, J.W., Wang, S., Yehuda, R., Riney, S., Charney, S., & Southwick, S.M. (2001). Psychogenic lowering of urinary cortisol levels linked to increased emotional numbing and a shame-depressive syndrome in combat-related posttraumatic stress disorder. *Psychosomatic Medicine*, 63, 387-401.
- Matthews, G. (1987) Personality and multidimensional arousal: A study of two dimensions of extraversion. *Personality and Individual Differences*, 8, 9-16.
- Matthews, G. (1996). Individual differences in driver stress and performance. *Proceedings of the Human Factors Ergonomics Society*, 40, 579-583.
- Matthews, G. (2001). A transactional model of driver stress. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 133-163). Mahwah, NJ: Erlbaum.
- Matthews, G. (2002). Towards a transactional ergonomics for driver stress and fatigue. *Theoretical Issues in Ergonomic Science*, 3, 195-211.
- Matthews, G., & Campbell, S.E. (1999). Individual differences in stress response and working memory. *Proceedings of the Human Factors and Ergonomics Society*, 43, 634-638.
- Matthews, G., Campbell, S. E., Desmond, P. A., Huggins, J., Falconer, S., & Joyner, A. (1999). Assessment of task-induced state change: Stress, fatigue and workload components. In M. Scerbo (Ed.), *Automation technology and human performance: Current research and trends* (pp. 199-203). Hillsdale, NJ: Erlbaum.
- Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., Grier, R., & Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, 2, 315-340.
- Matthews, G., & Davies, D. R. (1998). Arousal and vigilance: The role of task factors. In R. R. Hoffman, M. F. Sherrick, & J. S. Warm (Eds.), *Viewing psychology as a whole: The integrative science of William N. Dember* (pp. 113-144). Washington, DC: American Psychological Association.
- Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B. (2000). *Human performance: Cognition, stress and individual differences*. Hove, England: Psychology Press/Taylor & Francis.
- Matthews, G., & Desmond, P.A. (2002). Task-induced fatigue states and simulated driving performance. *Quarterly Journal of Experimental Psychology*, 55, 659-686.
- Matthews, G., Desmond, P. A., Joyner, L. A., & Carcary, B. (1997). A comprehensive questionnaire measure of driver stress and affect. In E. Carbonell-Vaya & J. A. Rothengatter (Eds.), *Traffic and transport psychology: Theory and application* (pp. 317-326). Amsterdam: Pergamon Press.



- Matthews, G., Dorn, L., Hoyes, T. W., Davies, D. R., Glendon, A. I., & Taylor, R. G. (1998). Driver stress and performance on a driving simulator. *Human Factors*, 40, 136-149.
- Matthews, G., Joyner, L., Gilliland, K., Campbell, S.E., Huggins, J., & Falconer, S. (1999). Validation of a comprehensive stress state questionnaire: Towards a state 'Big Three'? In I. Mervielde, I.J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe, Vol. 7* (pp. 335-350). Tilburg: Tilburg University Press.
- Matthews, G., & Margetts, I. (1991). Self-report arousal and divided attention: A study of performance operating characteristics. *Human Performance*, 4, 107-125.
- Matthews, G., Warm, J. S., Dember, W. N., Mizoguchi, H., & Smith, A. P. (2001). The common cold impairs visual attention, psychomotor performance and task engagement. *Proceedings of the Human Factors and Ergonomics Society*, 45, 1377-1381.
- May, P., Molloy, R., & Parasuraman, R. (1993, October). *Effects of automation reliability and failure rate on monitoring performance in a multitask environment*. Paper presented at the 39th Annual Meeting of the Human Factors Society, Seattle, WA.
- Mayleben, D.W. (1998). *Cerebral bloodflow velocity during sustained attention*. Unpublished doctoral dissertation, University of Cincinnati, Cincinnati, OH.
- Maxwell, S.E., & Delaney, H.D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Erlbaum.
- Mouloua, M., Gilson, R., & Hancock, P.A. (2003). Human-centered design of unmanned aerial vehicles. *Ergonomics and Design*, 11, 6-11.
- Norman, D.A., & Bobrow, D.G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44-64.
- Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97-113.
- Parasurman, R., & Riley, V. (1997). Humans and automation: use, misuse, disuse, abuse. *Human Factors*, 39, 230-253.
- Parasurman, R., Warm, J.S., & See, J.E. (1998). Brain systems of vigilance. In R. Parasurman (Ed.), *The attentive brain* (pp. 221-256). Cambridge, MA: MIT Press.
- Paus, T., Zatorre, R.J., Hofle, N., Caramanos, Z., Gotman, J., Petrides, M., & Evans, A. (1997). Time related changes in neural systems underlying attention and arousal during performance of an auditory vigilance task. *Journal of Cognitive Neuroscience*, 9, 392-408.



- Posner, M.I., & Petersen, S.E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25-42.
- Quinlan, P.T. (2003). Visual feature integration theory: Past, present, and future. *Psychological Bulletin*, 129, 643-673.
- Risberg, J. (1986). Regional cerebral blood flow in neuropsychology. *Neuropsychologia*, 24, 135-140.
- Reinartz, S. J., & Gruppe, T. R. (April, 1993). *Information requirements to support operator-automatic co-operation*. Paper presented at the Human Factors in Nuclear Safety Conference, London.
- Saucier, G., Goldberg, L.R. (2002). *Assessing the big five: Applications of 10 psychometric criteria to the development of marker scales*. Ashland, OH: Hogrefe & Huber.
- Schoenfeld, V.S., & Scerbo, M.W. (1997). Search differences for the presence and absence of features in sustained attention. *Proceedings of the Human Factors and Ergonomics Society*, 41, 1288-1292.
- Schoenfeld, V.S., & Scerbo, M.W. (1999). The effects of search differences for the presence and absence of features on vigilance performance and mental workload. In M.W. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends* (pp.177-182). Mahwah, NJ: Erlbaum.
- See, J.E., Howe, S.R., Warm, J.S., & Dember, W.N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*, 117, 230-249.
- Sheridan, D. (1980). Supervisory control. In G. Salvendy (Ed.) *Handbook of Human Factors* (pp. 1243-1268). New York: Wiley.
- Skosnik, P.D., Chatterton, R.T.Jr., Swisher, T., & Park, S. (2000). Modulation of attentional inhibition by norepinephrine and cortisol after psychological stress. *International Journal of Psychophysiology*, 36, 59-68.
- Sluiter, J.K., Frings-Dresen, M.H.W., van der Beek, A.J., & Meijman, T.F. (2001). The relation between work-induced neuroendocrine reactivity and recovery, subjective need for recovery, and health status. *Journal of Psychosomatic Research*, 50, 29-37.
- Stroobant, N., & Vingerhoets, G. (2000). Transcranial Doppler ultrasonography monitoring of cerebral hemodynamics during performance of cognitive tasks. *Neuropsychology Review*, 10, 213-231.
- Taylor, R.M. (1989). Situational awareness technique (SART): The development of a tool for



- aircrew systems design. *Proceedings of the NATO Advisory Group for Aerospace Research and Development (AGAR-CP-478)*.
- Thayer, R. (1989). *The biopsychology of mood and arousal*. New York: Oxford University Press.
- Treisman, A.M., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Bulletin*, 95, 15-48.
- U.S. Army Safety Center (n.d.). *Enhancing combat readiness through risk management: FY03 year end safety status*. Retrieved December 15, 2004, from <https://safety.army.mil/ipr/IPRFY03WEB2.ppt>
- Venables, L. & Fairclough, S.H. (2004). Establishing the psychophysiological variables that can identify and predict operator subjective state. *Proceedings of the Human Factors and Ergonomics Society*, 48, 90-94.
- Vingerhoets, G., & Stroobant, N. (1999). Lateralization of cerebral blood flow velocity changes during cognitive tasks. A simultaneous bilateral transcranial Doppler study. *Stroke*, 30, 2152-2158.
- Walker, G. H., Stanton, N. A., & Young, M. S. (2001) Where is computing driving cars? *International Journal of Human-Computer Interaction*, 13, 203-229.
- Warm, J.S. (1984). An introduction to vigilance. In J.S. Warm (Ed.), *Sustained attention in human performance* (pp.1-14). New York: Wiley.
- Warm, J.S. (1993). Vigilance in target detection. In B.M. Huey & C.D. Wickens (Eds.) *Workload Transition: Implications for individual and team performance* (pp. 139-170). Washington, DC: National Academy Press.
- Warm, J.S., & Dember, W.N. (1998). Tests of a vigilance taxonomy. In R.R. Hoffman, M.F. Sherrick, & J.S. Warm (Eds.), *Viewing psychology as a whole: The integrative science of William N. Dember* (pp. 87-112). Washington, DC: American Psychological Association.
- Warm, J.S., & Jerison, H.J. (1984). The psychophysics of vigilance. In J.S. Warm (Ed.) *Sustained Attention in Human Performance* (pp. 15-59). Chichester, UK: Wiley.
- Warm, J.S., Richter, D.O., Sprague, R.L., Porter, P.K., & Schumsky, D.A. (1980). Listening with a dual brain: hemispheric asymmetry in sustained attention. *Bulletin of the Psychonomic Society*, 15, 229-232.
- Washburn, D.A., Putney, R.T., & Tirre, W. (1998, November). *Attention profiles for speeded and unspeeded decision making*. Paper presented at the annual meeting of the Society for Judgment and Decision Making, Dallas, TX.



- Wiener, E. L. (1989). *Human factors of advanced technology ("glass cockpit") transport aircraft* (Rep. 177528). Moffett Field, CA: NASA Ames Research Center.
- Wickens, C.D., & Hollands, J.G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Wolf, O.T., Schommer, N.C., Hellhammer, D.H., McEwen, B.S., & Kirchbaum, C. (2001). The relationship between stress induced cortisol levels and memory differs between men and women. *Psychoneuroendocrinology*, 26, 711-720.
- Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors*, 44, 365-375.